

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 17, 2004, 06:17:00 ; Search time 11147.1 Seconds  
(without alignments)  
11447.105 Million cell updates/sec

Title: US-10-669-693-1

Perfect score: 2944

Sequence: 1 ttctctctgttgcttactc.....aaaaaaaaaaaaaaaaaaaaa 2944

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*

1: gb\_ba.\*

2: gb\_htg.\*

3: gb\_in.\*

4: gb\_om.\*

5: gb\_ov.\*

6: gb\_pat.\*

7: gb\_ph.\*

8: gb\_pl.\*

9: gb\_pr.\*

10: gb\_ro.\*

11: gb\_sts.\*

12: gb\_sy.\*

13: gb\_un.\*

14: gb\_vi.\*

15: em\_ba.\*

16: em\_fun.\*

17: em\_hum.\*

18: em\_in.\*

19: em\_mu.\*

20: em\_om.\*

21: em\_or.\*

22: em\_ov.\*

23: em\_pat.\*

24: em\_ph.\*

25: em\_pl.\*

26: em\_ro.\*

27: em\_sts.\*

28: em\_uni.\*

29: em\_vi.\*

30: em\_htg\_hum.\*

31: em\_htg\_inv.\*

32: em\_htg\_other.\*

33: em\_htg\_mus.\*

34: em\_htg\_pln.\*

35: em\_htg\_rod.\*

36: em\_htg\_mam.\*

37: em\_htg\_vrt.\*

38: em\_sy.\*

39: em\_htgo\_hum.\*

40: em\_htgo\_mus.\*

41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	2622.4	89.1	2626	9	AY358603	AY358603 Homo sapi
2	2621.4	89.0	2623	6	AX463930	AX463930 Sequence
3	2615.8	88.9	2619	9	BC033691	BC033691 Homo sapi
4	2612	88.7	2620	6	AX327331	AX327331 Sequence
5	2599.2	88.3	2604	6	AX430320	AX430320 Sequence
6	2596	88.2	2604	6	AX552230	AX552230 Sequence
7	2399.4	81.5	2513	6	AX876954	AX876954 Sequence
8	2399.4	81.5	2513	6	BD156406	BD156406 Primer fo
9	2399.4	81.5	2513	9	AK027605	AK027605 Homo sapi
10	2223.2	75.5	2261	6	AX780433	AX780433 Sequence
11	2223.2	75.5	2261	9	AF335278	AF335278 Homo sapi
12	1717.4	58.3	2114	6	AX552231	AX552231 Sequence
13	1513.4	51.4	1515	6	AX591678	AX591678 Sequence
14	1265.2	43.0	129402	9	AC011510	AC011510 Homo sapi
15	1263.6	42.9	173251	2	AC058798	AC058798 Homo sapi
16	885.6	30.1	6192	6	AX251439	AX251439 Sequence
17	658.4	22.4	2081	10	BC034202	BC034202 Mus muscu
18	613.2	20.8	6192	6	AX251440	AX251440 Sequence
19	601.2	20.4	812	6	AX381127	AX381127 Sequence
20	560.8	19.0	567	6	AR380149	AR380149 Sequence
21	556.8	18.9	562	6	AX779032	AX779032 Sequence
22	496.8	16.9	2065	4	RABCYCP2C	M20857 Rabbit cyto
23	495.2	16.8	2026	4	RABP450BX	L10912 Rabbit cyto
24	495.2	16.8	2028	4	S64259	S64259 cytochrome
25	490.4	16.7	2045	4	RABCYCP2B	M20856 Rabbit cyto
26	490.2	16.7	548	6	AX868269	AX868269 Sequence
27	490.2	16.7	548	6	BD148331	BD148331 Primer fo
28	489.2	16.6	1641	4	RABIIA10	L10236 Oryctolagus
29	488.6	16.6	1781	4	AB052255	AB052255 Sus scrof
30	485.6	16.5	2079	4	RABCYCF2A	M20855 Rabbit cyto
31	485.2	16.5	537	6	AX873328	AX873328 Sequence
32	485.2	16.5	537	6	BD153390	BD153390 Primer fo
33	482.8	16.4	2484	4	RABIIA11	L10237 Oryctolagus
34	482.4	16.4	1717	10	D86952	D86952 Mesocricetu
35	482.4	16.4	1802	4	RABP450CA	M18820 Rabbit cyto
36	481.2	16.3	2743	4	AB052256	AB052256 Sus scrof
37	475.2	16.1	1485	6	E10626	E10626 Human cDNA
38	475.2	16.1	1485	6	E10859	E10859 cDNA encodi
39	475.2	16.1	1485	6	E10860	E10860 cDNA encodi
40	475.2	16.1	1485	6	AR399327	AR399327 Sequence
41	475.2	16.1	1485	6	AR399328	AR399328 Sequence
42	475.2	16.1	1748	6	AX409692	AX409692 Sequence
43	475.2	16.1	1748	9	HUMCP1IA3A	M33318 Human cytoc
44	474.8	16.1	1485	9	AF209774	AF209774 Homo sapi
45	474.2	16.1	2004	6	AX477702	AX477702 Sequence

# ALIGNMENTS

RESULT 1  
AY358603  
LOCUS  
DEFINITION Homo sapiens clone DNA76395 Cytochrome p450 (UNQ891) mRNA, complete cds.  
ACCESSION AY358603  
VERSION AY358603.1 GI:37182327  
KEYWORDS FLI\_CDNA  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 2626)  
AUTHORS Clark,H.F., Gurney,A.L., Abaya,E., Baker,K., Baldwin,D., Brush,J.,  
Chen,J., Chow,B., Chui,C., Crowley,C., Deuel,B.,

Dowd,P., Eaton,D., Foster,J., Grimaldi,C., Gu,Q., Haas,P.E., Heldens,S., Huang,A., Kim,H.S., Klimowski,L., Jin,Y., Johnson,S., Lee,J., Lewis,L., Liao,D., Mark,M., Robbie,E., Sanchez,C., Schoenfeld,J., Seeshagiri,S., Simmons,L., Singh,J., Smith,V., Stinson,J., Vagts,A., Vandlen,R., Watanabe,C., Wieand,D., Woods,K., Xie,M.H., Yansura,D., Yi,S., Yu,G., Yuan,J., Zhang,M., Zhang,Z., Goddard,A., Wood,W.J. and Godowski,P.		The secreted protein discovery initiative (SPDI), a large-scale effort to identify novel human secreted and transmembrane proteins: A bioinformatics assessment	
Genome Res. 13 (10), 2265-2270 (2003)		12975309	
2 (bases 1 to 2626)		Clark,H.F.	
Direct Submission		Submitted (01-AUG-2003) Department of Bioinformatics, Genentech, Inc., 1 DNA Way, South San Francisco, CA 94080, USA	
JOURNAL		Location/Qualifiers	
FEATURES		1..2626	
source		/organism="Homo sapiens"	
		/mol_type="mRNA"	
		/db_xref="taxon:9606"	
		/clone="DNA76395"	
gene		1..2626	
CDS		44..1558	
		/locus_tag="UNQ891"	
		/locus_tag="UNQ891"	
		/note="PRO1906"	
		/codon_start=1	
		/product="Cytochrome p450"	
		/protein_id="AAQ8966.1"	
		/db_xref="GI:37182328"	
		/translation="MEATGFWALLLALLLALLLTLGSGTWARGLPPGTPPLILGNLIQLRPAALYSLMLRSLKXKGPVTIYILGPRPVVVLVGQEAIVEALGGQAEPSRGTVAMLEGTFDGHGVFFSNGERWRQRKFTMLADRLDGMKRGEBELIQAERCLIVETPQTEGRFPDPSLLAQATSNVVCLLGLRESYEDKSPQAVVRAAGTLLGLVSSQGGTTFMFSFPLRPGPHKQLLHVSTIAAFTVQVQHQGNLDSAPARDLIVAFLLMKAYEONPQTEFTNKNMLTVIYLLFAGTMTVSTVGYTLILLMKYTRFVQKWRBELNRELGAGQSLGDRITRLPYTDAVLHEAQRLLALVPMGIPRTLMLRTFRGYTLPGQTEVFLIGSLIHDNPIFKHPREFPNDFELDADGRFKHEAFLPESLGRVCLGELAKELFLFFTTILQAFSLSPCFPDFTLSLKPVTSLVGFNIPAPQLQVRFTDLHSTQTR"	
ORIGIN		Query Match 89.1%; Score 2622.4; DB 9; Length 2626;	
		Best Local Similarity 100.0%; Pred. No. 0;	
		Matches 2623; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	
QY	263	CGCGGAGCGCTGGGAGAGGAGAGGAGCGGACCTGCCAGATGGAGCGGACCGCACCT	322
DB	3	CGCGGAGCGCTGGGAGAGGAGAGGAGCGGACCTGCCAGATGGAGCGGACCGCACCT	62
QY	323	GGCGCTGCTGCTGGCGCTGGCGCTGCTCCTGCTGCTGACGCTGGCGCTGCCGGACCA	382
DB	63	GGCGCTGCTGCTGGCGCTGGCGCTGCTCCTGCTGCTGACGCTGGCGCTGCTGGGACCA	122
QY	383	GGCGCGAGGCCACTGCGCCCCCGGCGCCACGCCGCTACCACTGCTGGGAAACCTTCCTGC	442
DB	123	GGCGCGAGGCCACTGCGCCCCCGGCGCCACGCCGCTACCACTGCTGGGAAACCTTCCTGC	182
QY	443	AGCTACGGCCGGCGCTGTATTCAGGGCTCATCGGCTGAGTAAGAAGTACGACCGG	502
DB	183	AGCTACGGCCGGCGCGCTGTATTCAGGGCTCATCGGCTGAGTAAGAAGTACGACCGG	242
QY	503	TGTTACCACTACCTGGGACCTGGCGGCTGTGGTGCTGCTGGTGAGGAGGCTG	562
DB	243	TGTTACCACTACCTGGGACCTGGCGGCTGTGGTGCTGCTGGTGAGGAGGCTG	302
QY	563	TGCGGAGGCCCTGGGAGGCTCAGGCTTACGGCTTACGGCGGGGACCGCTAGCGATGC	622
DB	303	TGCGGAGGCCCTGGGAGGCTCAGGCTTACGGCTTACGGCGGGGACCGCTAGCGATGC	362
QY	623	TGGAAGGGAATTTTGATGGCCATGGGGTTTTCTTCTCCAAACGGGAGCGGTGGAGGCAGC	682



QY	1763	CCTTCCAGCTGCAAGTGCCTGCCACTGACCTTCACTCCACACGCGACGACGATGAAGGA	1822
Db	1503		
QY	1823	AGGCAACTTGAAGTGGTGGGTGCCAGGACGGTGCCTCCAGCTCTCAACAGTGGGCATGG	1882
Db	1563		
QY	1883	ACAGGTTAATGTCTCCAGAGTGTACTGCGAGCGACCAATTTACACGCCCTGCAGTTG	1942
Db	1623	ACAGGTTAATGTCTCCAGAGTGTACTGCGAGCGACCAATTTACACGCCCTGCAGTTG	1682
QY	1943	TTTTCCGGAGTCTGTCACCGGCCACACGCTCACTTGACTATGCTGCTTAAGATGCACA	2002
Db	1683		
QY	2003	ACCGCACACCCATACAACTATCAAGGGCCACAAGCAACTGCTGGGTAGCTTTCCACA	2062
Db	1743	ACCGCACACCCATACAACTATCAAGGGCCACAAGCAACTGCTGGGTAGCTTTCCACA	1802
QY	2063	GACATAAATATAGTCCATCTGCAATCAAAAGCAATAGCCAGGTAACCCACCACTCCCC	2122
Db	1803	GACATAAATATAGTCCATCTGCAATCAAAAGCAATAGCCAGGTAACCCACCACTCCCC	1862
QY	2123	TGATCTGCAGCCACACAGTGGGAGTCTGGCTGTCACITTCACAGGCCACAGAAAAGGCC	2182
Db	1863		
QY	2183	ACACATGTTCAACAGCTCACACGCCCTCTCCATTCATCGAACTTCTCAGTGTCCCTGTCCC	2242
Db	1923	ACACATGTTCAACAGCTCACACGCCCTCTCCATTCATCGAACTTCTCAGTGTCCCTGTCCC	1982
QY	2243	TGTTGCTTGCGCACAGGAAACAGATGCCCCCTCCGGGTCATGCCACCCACAGACTGTGCG	2302
Db	1983	TGTTGCTTGCGCACAGGAAACAGATGCCCCCTCCGGGTCATGCCACCCACAGACTGTGCG	2042
QY	2303	CTGTCTATGGCCCCCACTCATGTCTCCCTCTCTTTGGGTACACCACTCTCCAGGCTGTGAC	2362
Db	2043	CTGTCTATGGCCCCCACTCATGTCTCCCTCTCTTTGGGTACACCACTCTCCAGGCTGTGAC	2102
QY	2363	CACCGATGTCACACACCCCCAACCTTGTCCACACAGTACCCAGTACGACATCGTC	2422
Db	2103	CACCGATGTCACACACCCCCAACCTTGTCCACACAGTACCCAGTACGACATCGTC	2162
QY	2423	CTGGCTCCCCAGAGTATCTTCCCACTGAGACACGCGCGCCCCACAGAGGCACAGTCCCCA	2482
Db	2163	CTGGCTCCCCAGAGTATCTTCCCACTGAGACACGCGCGCCCCACAGAGGCACAGTCCCCA	2222
QY	2483	GCCACCTCTGCACTGCAGCCCTCAGTCACCCCTTTTAAAGACCCCTGATTTACCAAAAT	2542
Db	2223	GCCACCTCTGCACTGCAGCCCTCAGTCACCCCTTTTAAAGACCCCTGATTTACCAAAAT	2282
QY	2543	GCAAAACATCTGGGTCTGGATTATGCAACAGAGACTTTGGACATACAGGACCCCTCAGA	2602
Db	2283	GCAAAACATCTGGGTCTGGATTATGCAACAGAGACTTTGGACATACAGGACCCCTCAGA	2342
QY	2603	CCGGAGAAACACTGCCCCAACCCCAACACGTCCTTATGTAAACACGTGAAAAGCGGCCCC	2662
Db	2343	CCGGAGAAACACTGCCCCAACCCCAACACGTCCTTATGTAAACACGTGAAAAGCGGCCCC	2402
QY	2663	TGCTGCCCTTCCACACACATACACTCACTGATCTACAGCCCTGTTCGGGGTTCAGA	2722
Db	2403	TGCTGCCCTTCCACACACATACACTCACTGATCTACAGCCCTGTTCGGGGTTCAGA	2462
QY	2723	GTCCCCACTAGACCCAGTGGAGGGGTAGAGACCAAGTAGGGGCCAGTTTCCAATTCA	2782
Db	2463	GTCCCCACTAGACCCAGTGGAGGGGTAGAGACCAAGTAGGGGCCAGTTTCCAATTCA	2522
QY	2783	CCTGTGAGGAGTGAGCCGGATCTGACGTTTCTTGTGACTTAAGGGTCCGGCTTGGGAAT	2842
Db	2523	CCTGTGAGGAGTGAGCCGGATCTGACGTTTCTTGTGACTTAAGGGTCCGGCTTGGGAAT	2582

Qy	2843	TAAAGTTTGTTCTGGCCTTTAGCTTAAAAA	2886
Db	2583	TAAAGTTTGTTCTGGCCTTTAGCTTAAAAA	2626
RESULT 2			
AX463930			
LOCUS	AX463930	2623 bp	DNA linear PAT 16-JUL-2002
DEFINITION	Sequence 63 from Patent WO0140466.		
ACCESSION	AX463930		
VERSION	AX463930.1	GI:21898982	
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	Baker, K.P., Beresini, M., Deforge, L., Desnoyers, L., Filvaroff, E., Gao, W.Q., Gerritsen, M.E., Goddard, A., Godowski, P.J., Gurney, A.I., Sherwood, S., Smith, V., Stewart, T.A., Tumas, D., Watanabe, C.K., Wood, W.L. and Zhang, Z.		
TITLE	Secreted and transmembrane polypeptides and nucleic acids encoding same		
JOURNAL	Patent: WO 0140466-A 53 07-JUN-2001;		
FEATURES	Genentech Inc. (US)		
source	Location/Qualifiers		
	1. 2623		
	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:9606"		
ORIGIN			
	Query Match	89.0%;	Score 2621.4; DB 6; Length 2623;
	Best Local Similarity	100.0%;	Pred. No. 0;
	Matches 2622; Conservative	0;	Mismatches 1; Indels 0; Gaps 0;
Qy	264	GCGAGCGCTGGAGAGGAGAGGAGCGGACCTGCGAGATGAGGCGGACCTG 323	
Db	1	GCGAGCGCTGGAGAGGAGAGGAGCGGACCTGCGAGATGAGGCGGACCTG 60	
Qy	324	GGCGCTGCTGGCGTGGCGTGTCTCTGTCTGTGA CGCTGGCGTGTCCGGACCG 383	
Db	61	GGCGCTGCTGTGGCGTGGCGTGTCTCTGTCTGTGA CGCTGGCGTGTCCGGACCG 120	
Qy	384	GGCCCGAGGCGACCTGCCCCCGGGCCAGCGCGCTACCACTCTGGAAACCTCTGCA 443	
Db	121	GGCCCGAGGCGACCTGCCCCCGGGCCAGCGCGCTACCACTCTGGAAACCTCTGCA 180	
Qy	444	GTTACGCGCCGGCGCTGTATTACGGGCTCATGCGGCTGAGTAAAGTACGAGCCGT 503	
Db	181	GTTACGCGCCGGCGCTGTATTACGGGCTCATGCGGCTGAGTAAAGTACGAGCCGT 240	
Qy	504	GTTACGCGCTGGGACCTGGCGGCTGTGTGTCTGTGTGTGTGTGTGTGTGTGTGT 563	
Db	241	GTTACGCGCTGGGACCTGGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 300	
Qy	564	GCGGAGGCGCTGGGAGGTGAGCTGAGGAGTTACGCGCGCGGGAACGTAGCGATGCT 623	
Db	301	GCGGAGGCGCTGGGAGGTGAGCTGAGGAGTTACGCGCGCGGGAACGTAGCGATGCT 360	
Qy	624	GGAAGGACCTTTTGATGGGCATGGGCTTTTCTTCTCAACGGGGAGCGGTGGAGGCGCT 683	
Db	361	GGAAGGACCTTTTGATGGGCATGGGCTTTTCTTCTCAACGGGGAGCGGTGGAGGCGCT 420	
Qy	684	GAGGAAGTTTACCATGCTTGTCTTGGCGGACCTGGGCATGGGGAAGCGAAGGCGAGGA 743	
Db	421	GAGGAAGTTTACCATGCTTGTCTTGGCGGACCTGGGCATGGGGAAGCGAAGGCGAGGA 480	
Qy	744	GCTGATCCAGGCGGAGCGCGGTGTCTGTGTGAGACATTCAGGGGACAGAGGAGCGCC 803	
Db	481	GCTGATCCAGGCGGAGCGCGGTGTCTGTGTGAGACATTCAGGGGACAGAGGAGCGCC 540	
Qy	804	ATTGATCCCTCCCTGCTGCTGGCGGACCGCCACCTCCAACTGAGTCTGTCTCCCTCTTT 863	

Db 541 ATTGATCCCTCCCTGCTGCTGCGCCAGGCCACCTCCACAGTAGTCTGCTCCCTCCTCTT 600  
 QY 864 TGGCCCTCCGCTTCTCTATAGAGATAAGAGATTCCAGGCCGTGGTCCGGGACAGCTGGTGG 923  
 Db 601 TGGCCCTCCGCTTCTCTATAGAGATAAGAGATTCCAGGCCGTGGTCCGGGACAGCTGGTGG 660  
 QY 924 TACCTCTGTGGAGTCAAGTCCCGAGGGGGTCAAGACCTACAGATGTTCTCTGGTTCCT 983  
 Db 661 TACCTCTGTGGAGTCAAGTCCCGAGGGGGTCAAGACCTACAGATGTTCTCTGGTTCCT 720  
 QY 984 GGGGCCCTCCAGGCCCCCAAGCAGCTCTCTCCACACGCTCAGACCTTCGGCTGCCTT 1043  
 Db 721 GGGGCCCTCCAGGCCCCCAAGCAGCTCTCTCCACACGCTCAGACCTTCGGCTGCCTT 780  
 QY 1044 CACAGTCCGGCAGGTGACAGCACCCAGGGGACCTGATGTTTGGGGCCCGCAGCTGA 1103  
 Db 781 CACAGTCCGGCAGGTGACAGCACCCAGGGGACCTGATGTTTGGGGCCCGCAGCTGA 840  
 QY 1104 CTTGTGATGCTTCTCTCTGAAAGATGGACAGGAGGAAACAAACCCAGGACAGATTT 1163  
 Db 841 CTTGTGATGCTTCTCTCTGAAAGATGGACAGGAGGAAACAAACCCAGGACAGATTT 900  
 QY 1164 CACCAACAGAACATGCTGATGACAGTCAATTTATTTGCTGTTTCTGGACGATGACGCT 1223  
 Db 901 CACCAACAGAACATGCTGATGACAGTCAATTTATTTGCTGTTTCTGGACGATGACGCT 960  
 QY 1224 CAGCACCAAGCTCGGCTATACCTCTCTCTGATGAATACCTCATGTCACAAAGTG 1283  
 Db 961 CAGCACCAAGCTCGGCTATACCTCTCTCTGATGAATACCTCATGTCACAAAGTG 1020  
 QY 1284 GGTAAGTGAAGAGCTGAATGCGGAGCTGGGGGTGGCCAGGACCAAGCTTAGGGACCG 1343  
 Db 1021 GGTAAGTGAAGAGCTGAATGCGGAGCTGGGGGTGGCCAGGACCAAGCTTAGGGACCG 1080  
 QY 1344 TACCGGCTCCCTTACACGAGCGGTCTGATGAGGCGGAGCGGCTGCGCGTGGT 1403  
 Db 1081 TACCGGCTCCCTTACACGAGCGGTCTGATGAGGCGGAGCGGCTGCGCGTGGT 1140  
 QY 1404 GCCCATGGGAATACCCCGCACCTCTATGCGGACCAACCCGCTTCCGAGGGTACACCTGCG 1463  
 Db 1141 GCCCATGGGAATACCCCGCACCTCTATGCGGACCAACCCGCTTCCGAGGGTACACCTGCG 1200  
 QY 1464 CAGGGGACAGGAGTCTTCCCTCTCTGCTCCATCTGATGACACCCCAACATCTCAA 1523  
 Db 1201 CAGGGGACAGGAGTCTTCCCTCTCTGCTCCATCTGATGACACCCCAACATCTCAA 1260  
 QY 1524 GCACCCAGAGAGTTCAACCCAGACCGTTTCTGATGAGATGGAGCGGTTCAAGGAAGA 1583  
 Db 1261 GCACCCAGAGAGTTCAACCCAGACCGTTTCTGATGAGATGGAGCGGTTCAAGGAAGA 1320  
 QY 1584 TGAGGCGTTCTGCTTCTCTTATAGGAAGCGTGTGCTTGGAGAGGCGCTGGCAAA 1643  
 Db 1321 TGAGGCGTTCTGCTTCTCTTATAGGAAGCGTGTGCTTGGAGAGGCGCTGGCAAA 1380  
 QY 1644 AGCGAGCTTCTCTTCTTACACCATCTTACAGGCTTCTTCCCTGGAGAGCGGCTG 1703  
 Db 1381 AGCGAGCTTCTCTTCTTACACCATCTTACAGGCTTCTTCCCTGGAGAGCGGCTG 1440  
 QY 1704 CCGCGGACACCTCAGCTCAAGCCCAAGCGTCTGATGAGATGGAGCGGTTCAACATTTCCCGCAGC 1763  
 Db 1441 CCGCGGACACCTCAGCTCAAGCCCAAGCGTCTGATGAGATGGAGCGGTTCAACATTTCCCGCAGC 1500  
 QY 1764 CTTCCAGCTGCAAGTCCGTCCTTCTTATAGGAAGCGTGTGCTTGGAGAGGCGCTGGCAAA 1823  
 Db 1501 CTTCCAGCTGCAAGTCCGTCCTTCTTATAGGAAGCGTGTGCTTGGAGAGGCGCTGGCAAA 1560  
 QY 1824 GGCACCTTGAAGTGGTGGTGGCCAGGAGCGTCTCAGGCTCAACAGTGGGATGGA 1883  
 Db 1561 GGCACCTTGAAGTGGTGGTGGCCAGGAGCGTCTCAGGCTCAACAGTGGGATGGA 1620  
 QY 1884 CAGGGTTAATGCTCCAGAGTGTACATGTCAGGAGCGCACATTTACACGCTCGAGTTGT 1943

Db 1621 CAGGGTTAATGCTCCAGAGTGTACACTGAGGAGCCACATTTACACGCTGCGAGTTGT 1680  
 QY 1944 TTTCCGAGTCTGTCCCAAGCCACAGCTCAGTCTGACTCATGCTGCTAGATGACAAA 2003  
 Db 1681 TTTCCGAGTCTGTCCCAAGCCACAGCTCAGTCTGACTCATGCTGCTAGATGACAAA 1740  
 QY 2004 CGGCACACCCATACACAACTCAAGGGCCACAAAGCAACTGCTGGTGTAGCTTTCCACAG 2063  
 Db 1741 CGGCACACCCATACACAACTCAAGGGCCACAAAGCAACTGCTGGTGTAGCTTTCCACAG 1800  
 QY 2064 ACATAAATATAGTCCATCTGCAATCAAGCAATAGCAGGTAACCCACCACTCCCT 2123  
 Db 1801 ACATAAATATAGTCCATCTGCAATCAAGCAATAGCAGGTAACCCACCACTCCCT 1860  
 QY 2124 GGATCTGAGCCCAAGCTGGAGTCTGGTGTCACTTCAAGCCACAGAAAGCGCA 2183  
 Db 1861 GGATCTGAGCCCAAGCTGGAGTCTGGTGTCACTTCAAGCCACAGAAAGCGCA 1920  
 QY 2184 CACATGTTTCAAGCTCACAAGCCCTCTCCATTCATCGAACTTCTCAGTGTCCCTGCTCCCT 2243  
 Db 1921 CACATGTTTCAAGCTCACAAGCCCTCTCCATTCATCGAACTTCTCAGTGTCCCTGCTCCCT 1980  
 QY 2244 GGTGCTTGSCACAGGAAACAGATGCCCTCCGGGTCAATGCCACCCAGAGACTGTGCG 2303  
 Db 1981 GGTGCTTGSCACAGGAAACAGATGCCCTCCGGGTCAATGCCACCCAGAGACTGTGCG 2040  
 QY 2304 TGTCTATGGCCCAACTCATGCTCTCTTGGCTTACACACTCTCCAGCTGTGACC 2363  
 Db 2041 TGTCTATGGCCCAACTCATGCTCTCTTGGGTACACACTCTCCAGCTGTGACC 2100  
 QY 2364 ACCGATGTCACACACCCCAACCACTTGTCCACACAGTACCCAGTACGACATCGTCC 2423  
 Db 2101 ACCGATGTCACACACCCCAACCACTTGTCCACACAGTACCCAGTACGACATCGTCC 2160  
 QY 2424 TGGCTCCCAGAGTATCTTCCCACTGAGACAAGCGCGCCCAAGAGGACAGTGTGCG 2483  
 Db 2161 TGGCTCCCAGAGTATCTTCCCACTGAGACAAGCGCGCCCAAGAGGACAGTGTGCG 2220  
 QY 2484 CCACCTCTGCAACTGCGGCTCAGTCAACCCCTTTTAAAGCAGCTGATTTACCAAAATG 2543  
 Db 2221 CCACCTCTGCAACTGCGGCTCAGTCAACCCCTTTTAAAGCAGCTGATTTACCAAAATG 2280  
 QY 2544 CAAACACATCTGGTCTGCGATTATGCAAGAGACTTTGGACATAGAGAGCCCTCAGAC 2603  
 Db 2281 CAAACACATCTGGTCTGCGATTATGCAAGAGACTTTGGACATAGAGAGCCCTCAGAC 2340  
 QY 2604 CGGAGAAACCTGCGCAACCCCAACACGCTGTTATGTAACACGTTGGAAGCGGCCCT 2663  
 Db 2341 CGGAGAAACCTGCGCAACCCCAACACGCTGTTATGTAACACGTTGGAAGCGGCCCT 2400  
 QY 2664 GCTGCCCTCCACACACATACACTCACTGATCTACAGCCCTGTTCCGCGTCCAGAG 2723  
 Db 2401 GCTGCCCTCCACACACATACACTCACTGATCTACAGCCCTGTTCCGCGTCCAGAG 2460  
 QY 2724 TCCCACTAGACCCAGTGAAGGGTTAGAGACCAAGTAGGGCCAGTTTCCAATTCACC 2783  
 Db 2461 TCCCACTAGACCCAGTGAAGGGTTAGAGACCAAGTAGGGCCAGTTTCCAATTCACC 2520  
 QY 2784 CTGTAGGAGTGAAGCCGATCTGAGTCTTGTGATCTTAAGGGTCCGGCTTGGGAAT 2843  
 Db 2521 CTGTAGGAGTGAAGCCGATCTGAGTCTTGTGATCTTAAGGGTCCGGCTTGGGAAT 2580  
 QY 2844 AAGATTTGTTCTGCGCTTTAGCTAAAAAATAAAAAA 2886  
 Db 2581 AAGATTTGTTCTGCGCTTTAGCTAAAAAATAAAAAA 2623

RESULT 3  
 BC033691  
 LOCUS  
 DEFINITION  
 ACCESSION

BC033691 2619 bp mRNA linear PRI 06-OCT-2003  
 Homo sapiens cytochrome P450, family 2, subfamily S, polypeptide 1,  
 mRNA (cDNA clone MGC:44853 IMAGE:5212609), complete cds.  
 BC033691

```
BC033691.1 GI:21707057
MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, P.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Haie, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Schweetz, T.E., Brownstein, M.J., Uesdin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raja, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smallos, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
22388257
PUBMED 12477932
REFERENCE
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (02-JUL-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK
COMMENT NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: gcapsb@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: http://www.nisc.nih.gov/
Contact: nisc_mgc@hgrl.nih.gov
Akhtar, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,
Maduro, Q.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
McDowell, J., Pearson, R., Stantropop, S., Thomas, P.J., Touchman, J.W.,
Thurgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
Young, A., Zhang, L.-H. and Green, E.D.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAC Plate: 68 Row: 0 Column: 11
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 20522238.
FEATURES
SOURCE
1. .2619
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:44853 IMAGE:5212609"
/tissue_type="Blood, adult leukocytes"
/clone_lib="NIH_MGC_118"
/lab_host="DH10B"
/notes="Vector: pCMV-SPORT6"
1. .2619
gene
```

Db	601	ATAAGGAGTTCCAGGCGGTGGTCCGGGCAAGCTGGTGGTACCCCTGCTGGGAGTCAGTCTCC	660
Qy	947	AGGGGGGTACAGACCTACGAGATGTTCTCTGGTTCTCTCGCGCCCTGCCAGGCCCCACA	1006
Db	661	AGGGGGGTACAGACCTACGAGATGTTCTCTGGTTCTCTCGCGCCCTGCCAGGCCCCACA	720
Qy	1007	AGCAGCTCTCCACACGTACAGACCTTGGCTGCTTTCACAGTCCGGCAGGTGAGCAGC	1066
Db	721	AGCAGCTCTCCACACGTACAGACCTTGGCTGCTTTCACAGTCCGGCAGGTGAGCAGC	780
Qy	1067	ACRAGGGGAACCTCGATGCTTTGGGCCCCGACAGTGACCTTGTGATGCTTCCCTGCTGA	1126
Db	781	ACRAGGGGAACCTCGATGCTTTGGGCCCCGACAGTGACCTTGTGATGCTTCCCTGCTGA	840
Qy	1127	AGATGGCACAGGAGGAAACAAACCCAGGCACAGAAATTCACCAACAAGAACATGCTGATGA	1186
Db	841	AGATGGCACAGGAGGAAACAAACCCAGGCACAGAAATTCACCAACAAGAACATGCTGATGA	900
Qy	1187	CAGTCATTTATTTCTGTTTCTGGGACGATGACGTTACAGCCACAGTCCGGCTATATACC	1246
Db	901	CAGTCATTTATTTCTGTTTCTGGGACGATGACGTTACAGCCACAGTCCGGCTATATACC	960
Qy	1247	TCCTGCTCCTGATGAAATACCTCATGTCCAAAAGTGGTACGTGAGGAGCTGAAATCGGG	1306
Db	961	TCCTGCTCCTGATGAAATACCTCATGTCCAAAAGTGGTACGTGAGGAGCTGAAATCGGG	1020
Qy	1307	AGCTGGGGGCTGGCAGGCACCAAGCTTAGGGGACCGTACCGGCTCTCCCTTACACGAGC	1366
Db	1021	AGCTGGGGGCTGGCAGGCACCAAGCTTAGGGGACCGTACCGGCTCTCCCTTACACGAGC	1080
Qy	1367	CGGTTCTCATGAGCGCAGCGGCTGCTGGCGCTGGTGCCCATGGGAATACCCCGACCC	1426
Db	1081	CGGTTCTCATGAGCGCAGCGGCTGCTGGCGCTGGTGCCCATGGGAATACCCCGACCC	1140
Qy	1427	TCATGGGAAACAACGCTTCCGAGGGTACACCTTGCCCCACAGGCACGGAGGTCTTCCCC	1486
Db	1141	TCATGGGAAACAACGCTTCCGAGGGTACACCTTGCCCCACAGGCACGGAGGTCTTCCCC	1200
Qy	1487	TCCTTGGCTCCATCTGTCATGACCCCAACATCTTCAAGCACCCAGAAGAGTTCAAACCCAG	1546
Db	1201	TCCTTGGCTCCATCTGTCATGACCCCAACATCTTCAAGCACCCAGAAGAGTTCAAACCCAG	1260
Qy	1547	ACCGTTTCTCGATGACAGATGGACGGTTTCAGGAAGCATGAGGGCTTCCCTGCCCTTTCCT	1606
Db	1261	ACCGTTTCTCGATGACAGATGGACGGTTTCAGGAAGCATGAGGGCTTCCCTGCCCTTTCCT	1320
Qy	1607	TAGGGAGCGTGTCTGCTTGGAGGGCTGGCAAGCGAGCTCTTCTCTCTCTTCA	1666
Db	1321	TAGGGAGCGTGTGTGCTTGGAGGGCTGGCAAGCGAGCTCTTCTCTCTCTTCTCA	1380
Qy	1667	CCACCATCTCAAGCCTTCTCCCTGGAGAGCCGTCGCCGCCGGAACACCTTGAGCCTCA	1726
Db	1381	CCACCATCTCAAGCCTTCTCCCTGGAGAGCCGTCGCCGCCGGAACACCTTGAGCCTCA	1440
Qy	1727	AGCCCCCGTCAGTGGCCCTTTTCAACATTCGCCACGCTTCCAGCTGCAAGTCGGTCCA	1786
Db	1441	AGCCCCCGTCAGTGGCCCTTTTCAACATTCGCCACGCTTCCAGCTGCAAGTCGGTCCA	1500
Qy	1787	CTGACCTTTCACCTCACACGACGACAGATGAAGAGGCAACCTTGGAAAGTGGTGGGTGC	1846
Db	1501	CTGACCTTTCACCTCACACGACGACAGATGAAGAGGCAACCTTGGAAAGTGGTGGGTGC	1560
Qy	1847	CCAGGACGGTGTCTCCAGCCTCAACAGTGGGCATGGACAGGGTTAAATGTCTCCAGAGTGT	1906
Db	1561	CCGGGACGGTGTCTCCAGCCTCAACAGTGGGCATGGACAGGGTTAAATGTCTCCAGAGTGT	1620
Qy	1907	ACACTGACGAGGACCAATTTACAGCCTGCAAGTTGTTTTCCGAGTCTGTGCCACCGCC	1966
Db	1621	ACACTGACGAGGACCAATTTACAGCCTGCAAGTTGTTTTCCGAGTCTGTGCCACCGCC	1680
Qy	1967	CACACGCTCACTTGATCATGCTGTAAAGATGCAACCCGACACCCCATACAACTATCA	2026
Db	1681	CACACGCTCACTTGATCATGCTGTAAAGATGCAACCCGACACCCCATACAACTATCA	1740

QY	2027	AGGCCACAAGCAACTGCTGGGTAGCTTTTCACAGACATAAATATAGTTCGATCTGC	2086
DB	1741	AGGCCACAAGCAACTGCTGGGTAGCTTTTCACAGACATAAATATAGTTCGATCTGC	1800
QY	2087	TCACAAGCATATAGCCAGGTAAACCCACCAACTCCCTGGATCTGCAGCCACACGCTGGGA	2146
DB	1801	TCACAAGCATATAGCCAGGTAAACCCACCAACTCCCTGGATCTGCAGCCACACGCTGGGA	1860
QY	2147	GTCTGGCTGTCACTTCACAAGCCACAGAAACGGCCACATATGTTACAGCTTCACAGCC	2206
DB	1861	GTCTGGCTGTCACTTCACAAGCCACAGAAACGGCCACATATGTTACAGCTTCACAGCC	1920
QY	2207	CTCTCCATTCAATCGAACTTCTCAGTGTCCCTGTCCCTGTGCTGTGCGCACAGGAAACAGCA	2266
DB	1921	CTCTCCATTCAATCGAACTTCTCAGTGTCCCTGTCCCTGTGCTGTGCGCACAGGAAACAGCA	1980
QY	2267	TGCCCCCTCGGGGTATGCCACCCAGAGACTGTGCTGTCTATGGCCCCCAACTCATGCT	2326
DB	1981	TGCCCCCTCGGGGTATGCCACCCAGAGACTGTGCTGTCTATGGCCCCCAACTCATGCT	2040
QY	2327	CCCTCTCTTGGCTACACCACTCTCTCCAGCCTGTGACCAACCGATGTCCACACACCCCAAC	2386
DB	2041	CCCTCTCTTGGCTACACCACTCTCTCCAGCCTGTGACCAACCGATGTCCACACACCCCAAC	2100
QY	2387	CACCTGTGCCACAGCTACCCAGTAGACATGTCCTGTGCTGCCAGAGTATCTTCCCA	2446
DB	2101	CACCTGTGCCACAGCTACCCAGTAGACATGTCCTGTGCTGCCAGAGTATCTTCCCA	2160
QY	2447	CTGAGACACGCCGCCCCACAGAGGACAGTCCCGACCACTCTGCAACTGACAGCCCTC	2506
DB	2161	CTGAGACACGCCGCCCCACAGAGGACAGTCCCGACCACTCTGCAACTGACAGCCCTC	2220
QY	2507	AGTCACCCCTTTTAAAGCACCCCTGATTCTACCAAAATGCAAAACATCTGGGTCTGCGATT	2566
DB	2221	AGTCACCCCTTTTAAAGCACCCCTGATTCTACCAAAATGCAAAACATCTGGGTCTGCGATT	2280
QY	2567	ATGCACAGAGACTTTGACATACAGAGACCCCTCAGCCGAGGAGACACCTGCCCAACCCC	2626
DB	2281	ATGCACAGAGACTTTGACATACAGAGACCCCTCAGCCGAGGAGACACCTGCCCAACCCC	2340
QY	2627	AACACGTCTTATGTAAACCACTGTGAAAGGGGCCCTGTCTGCCCTCCACACACATAC	2686
DB	2341	AACACGTCTTATGTAAACCACTGTGAAAGGGGCCCTGTCTGCCCTCCACACACATAC	2400
QY	2687	ACATCTCATGATCTACAGCCCTGTTTCGGGTCTAGATCCCGACTAGACCCAGTGGAGG	2746
DB	2401	ACATCTCATGATCTACAGCCCTGTTTCGGGTCTAGATCCCGACTAGACCCAGTGGAGG	2460
QY	2747	GGTTAGAGACCAAGTAGGGCCAGTTTCCAAATTCACCTCTGTAGGAGTGTAGCCCGATCT	2806
DB	2461	GGTTAGAGACCAAGTAGGGCCAGTTTCCAAATTCACCTCTGTAGGAGTGTAGCCCGATCT	2520
QY	2807	GACGTCTCTTGTGACTTAAGGGTCCGGCTTGGGAATTAAGTTTGTCTTCTGGCCTTTAGC	2866
DB	2521	GACGTCTCTTGTGACTTAAGGGTCCGGCTTGGGAATTAAGTTTGTCTTCTGGCCTTTAGC	2580
QY	2867	CTAAAAA	2905
DB	2581	CTAAAAA	2619
RESULT 4			
AX327331			
LOCUS			
DEFINITION			
Sequence 16 from Patent WO0179468.			
ACCESSION			
AX327331			
VERSION			
AX327331.1			
GI:18097877			
KEYWORDS			
Homo sapiens (human)			
SOURCE			
Homo sapiens			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
PAT 07-JAN-2002			

REFERENCE  
1 Policky,J.L., Hafalia,A., Burford,N., Ring,H.Z., Lal,P.,  
AUTHORS Tribouley,C.M., Yao,M.G., Yue,H., Tang,F.T., Patterson,C., Das,D.,  
Sanjanwala,M.S., Gandhi,A.R., Reddy,R., Khan,F.A., Baughn,M.R.,  
Ramkumar,J., Griffin,J.A. and Au-Young,J.  
TITLE Drug metabolizing enzymes  
JOURNAL Patent: WO 0179468-A 16 25-OCT-2001;  
Incyte Genomics, Inc. (US)  
FEATURES Location/Qualifiers  
source 1..2620  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
/note="Incyte ID No: 1272843CB1"  
ORIGIN  
Query Match 88.7%; Score 2612; DB 6; Length 2620;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2612; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 272 CCTGGGAGAGGAGGAGCGGACCTGCGGAGATGGAGGCGACCGGACCTGGGCGGTGC 331  
DB 9 CCTGGGAGAGGAGGAGCGGACCTGCGGAGATGGAGGCGGACCGGACCTGGGCGGTGC 68  
QY 332 TGTGGCGCTGGCGCTGCTCTGCTGTGCTGACGCTGGCGCTGTGCGGACGAGGCGCGAG 391  
DB 69 TGTGGCGCTGGCGCTGCTCTGCTGTGCTGACGCTGGCGCTGTGCGGACGAGGCGCGAG 128  
QY 392 GCCACCTGCCCGCGGCGCCACCGCGTACCACTGCTGGGAAACCTCTGCGAGTACGGC 451  
DB 129 GCCACCTGCCCGCGGCGCCACCGCGTACCACTGCTGGGAAACCTCTGCGAGTACGGC 188  
QY 452 CCGGGCGCTATTACGGGCTCATGGGCTGAGTGAAGTAGCGGACCGGTGTTCACCA 511  
DB 189 CCGGGCGCTGTATTACGGGCTCATGGGCTGAGTGAAGTAGCGGACCGGTGTTCACCA 248  
QY 512 TCTACCTGGGACCTGGCGGCTGTGTGTGCTGTGGTGGGACGAGGCTGTGGGGAGG 571  
DB 249 TCTACCTGGGACCTGGCGGCTGTGTGTGCTGTGGTGGGACGAGGCTGTGGGGAGG 308  
QY 572 CCTTGGGAGGTGAGGCTGAGGAGTTACGCGCGCGGGGAAACCGTAGCGATGCTGGAAGGGA 631  
DB 309 CCTTGGGAGGTGAGGCTGAGGAGTTACGCGCGCGGGGAAACCGTAGCGATGCTGGAAGGGA 368  
QY 632 CTTTGTGATGCGCATGGGGTTTCTTCTCCAAAGGGGACGCGTGGAGGAGCTGAGGAAGT 691  
DB 369 CTTTGTGATGCGCATGGGGTTTCTTCTCCAAAGGGGACGCGTGGAGGAGCTGAGGAAGT 428  
QY 692 TTACCATGCTTGTCTGCGGACCTGGGCAATGGGCAAGCGAAGGCGAGGAGCTGATCC 751  
DB 429 TTACCATGCTTGTCTGCGGACCTGGGCAATGGGCAAGCGAAGGCGAGGAGCTGATCC 488  
QY 752 AGGCGGAGGCGCGGTGTCTGGTGGAGACATTCAGAGGGGACAGAGGACGCCATTGCGATC 811  
DB 489 AGGCGGAGGCGCGGTGTCTGGTGGAGACATTCAGAGGGGACAGAGGACGCCATTGCGATC 548  
QY 812 CCTCCCTGCTGCTGGCGGACGACCTCCACAGTAGTCTGCTCCTCTCTTTGGGCTCC 871  
DB 549 CCTCCCTGCTGCTGGCGGACGACCTCCACAGTAGTCTGCTCCTCTCTTTGGGCTCC 608  
QY 872 GCTTCTCTATGAGGATAAGGAGTTCCAGGCGGTGGTCCGGGACGCTGGTGTACCTGTC 931  
DB 609 GCTTCTCTATGAGGATAAGGAGTTCCAGGCGGTGGTCCGGGACGCTGGTGTACCTGTC 668  
QY 932 TGGGAGTCAGTCCAGGGGGGTGAGACCTACGAGATGTTCTCTGGTTCCTGCGGCCCC 991  
DB 669 TGGGAGTCAGTCCAGGGGGGTGAGACCTACGAGATGTTCTCTGGTTCCTGCGGCCCC 728  
QY 992 TGCAGGCGCCCAAGAGGCTCCTCCACAGCTGACACCTTGGCTGCGCTTACAGTCC 1051  
DB 729 TGCAGGCGCCCAAGAGGCTCCTCCACAGCTGACACCTTGGCTGCGCTTACAGTCC 788  
QY 1052 GGCAGGTGCAGCAGCACCAGGGGAACCTGGATGCTTCGGGCGCCGCGACGTGACCTTGTGC 1111

DB 789 GGCAGGTGCAGCAGCACCAGGGGAACCTGGATGCTTCGGGCCCCGCGACGTGACCTTGTGC 848  
QY 1112 ATGCTTCTCTGCTGAAGATGGCAGAGGAGAAACCCAGGCGACAGAAATTCACCAACA 1171  
DB 849 ATGCTTCTCTGCTGAAGATGGCAGAGGAGAAACCCAGGCGACAGAAATTCACCAACA 908  
QY 1172 AGAATGCTGATGACAGTCAATTTATTTGCTGTTTGTCTGGACGATGACGCTGAGCACA 1231  
DB 909 AGAATGCTGATGACAGTCAATTTATTTGCTGTTTGTCTGGACGATGACGCTGAGCACA 968  
QY 1232 CGGTGCGCTATACCTCTCTGCTCTGATGAAATACCTCTCATGTCCAAAAGTGGGTACGTG 1291  
DB 969 CGGTGCGCTATACCTCTCTGCTCTGATGAATACCTCTCATGTCCAAAAGTGGGTACGTG 1028  
QY 1292 AGGAGCTGAATCGGAGCTGGGGCTGGCCAGGACCAAGGCTTAGGGGACCGTACCCGCC 1351  
DB 1029 AGGAGCTGAATCGGAGCTGGGGCTGGCCAGGACCAAGGCTTAGGGGACCGTACCCGCC 1088  
QY 1352 TCCCTTACCGGAGCGGTTCTGCATGAGGGGAGGGCTGCTGGCGCTGTTGCCCATGG 1411  
DB 1089 TCCCTTACCGGAGCGGTTCTGCATGAGGGGAGCGGCTGCTGGCGCTGTTGCCCATGG 1148  
QY 1412 GAATACCCCGCACCTCATGCGGACCAACCGCTTCCGAGGGTACACCTGCCCCAGGGCA 1471  
DB 1149 GAATACCCCGCACCTCATGCGGACCAACCGCTTCCGAGGGTACACCTGCCCCAGGGCA 1208  
QY 1472 CGGAGGTCTTCCCTCTCTTGGCTCCATCTGTCATGACCCCAACATCTTCAAGCACCCAG 1531  
DB 1209 CGGAGGTCTTCCCTCTCTTGGCTCCATCTGTCATGACCCCAACATCTTCAAGCACCCAG 1268  
QY 1532 AAGAGTTCAACCCAGACCGTTTCTGATGACAGATGACCGTTTCAAGAGCATGAGGCGT 1591  
DB 1269 AAGAGTTCAACCCAGACCGTTTCTGATGACAGATGACCGTTTCAAGAGCATGAGGCGT 1328  
QY 1592 TCTGTCCTTCTCTTAGGGAAGCGTCTGCTCTTGGAGAGGGCTTGGCAAAAGCGGAGC 1651  
DB 1329 TCTGTCCTTCTCTTAGGGAAGCGTCTGCTCTTGGAGAGGGCTTGGCAAAAGCGGAGC 1388  
QY 1652 TCTTCTCTTCTTCCACCACCATCTTCAAGCGCTTCTCCCTGGAGAGCGCTGCGCGCGG 1711  
DB 1389 TCTTCTCTTCTTCCACCACCATCTTCAAGCGCTTCTCCCTGGAGAGCGCTGCGCGCGG 1448  
QY 1712 ACACCTTGAGCTTCAAGCCACCGTCTGATGCGCTTTTCAACATTTCCCGCAGCGCTTCAGC 1771  
DB 1449 ACACCTTGAGCTTCAAGCCACCGTCTGATGCGCTTTTCAACATTTCCCGCAGCGCTTCAGC 1508  
QY 1772 TGCAAGTCCGTCCCACTGACCTTCACTCCACCGCAGACCCAGATGAAGAGGCAACTT 1831  
DB 1509 TGCAAGTCCGTCCCACTGACCTTCACTCCACCGCAGACCCAGATGAAGAGGCAACTT 1568  
QY 1832 GGAAGTGGTGGGTGCCCAGGACCGTCTCCAGCTCAACAGTGGGCGATGGACAGGGTTA 1891  
DB 1569 GGAAGTGGTGGGTGCCCAGGACCGTCTCCAGCTCAAAGTGGGCGATGGACAGGGTTA 1628  
QY 1892 ATGCTCCAGAGTGTACACTGCGAGGAGCCACATTTACGCGCTGAGTGTGTTTCCGGA 1951  
DB 1629 ATGCTCCAGAGTGTACACTGCGAGGAGCCACATTTACGCGCTGAGTGTGTTTCCGGA 1688  
QY 1952 GTCTGTCCCGCGCCACAGCTCACTTGAATGCTGCTTAAGATGCACACCGCACAC 2011  
DB 1689 GTCTGTCCCGCGCCACAGCTCACTTGAATGCTGCTTAAGATGCACACCGCACAC 1748  
QY 2012 CCATACAACTACAAGGGCCACAAAGCAACTGCTGGGTTAGCTTTTCCACAGACATAAAT 2071  
DB 1749 CCATACAACTACAAGGGCCACAAAGCAACTGCTGGGTTAGCTTTTCCACAGACATAAAT 1808  
QY 2072 ATAGTCCATCTGCAATCAAGCAGATAGCCAGGTAAACCAACCACTCCCTGATCTGC 2131  
DB 1809 ATAGTCCATCTGCAATCAAGCAGATAGCCAGGTAAACCAACCACTCCCTGATCTGC 1868  
QY 2132 AGCCCAACAGTGGGAGTCTGCTGCTTCAAGGCGACAGAAAGCGGCGCACATGTT 2191





Qy	1285	GTACGTGAGGAGCTGAATCCGGAGCTGGGGCTGGCCAGGCACCAAGCCTTAGGGGACCGT	1344
Db	1021		
		GTACGTGAGGAGCTGAATCGGGAGCTGGGGCTGGCCAGGCACCAAGCCTTAGGGGACCGT	1080
Qy	1345	ACCCGCTCCCTTACACCGAGCGCGTCTCTGATCAGGCGCAGCGGCTGCTGGCGCTGGTG	1404
Db	1081		
		ACCCGCTCCCTTACACCGAGCGCGTCTCTGATCAGGCGCAGCGGCTGCTGGCGCTGGTG	1140
Qy	1405	CCCATGGGAATACCCCGCACCTCATGCGGACCAACCCGCTTCCGAGGGTACACCTCGCC	1464
Db	1141		
		CCCATGGGAATACCCCGCACCTCATGCGGACCAACCCGCTTCCGAGGGTACACCTCGCC	1200
Qy	1465	CAGGGCACGAGGCTTCCCGCTCCTTGGCTCCATCTCTGATGACCCCAACATCTTCAAG	1524
Db	1201		
		CAGGGCACGAGGCTTCCCGCTCCTTGGCTCCATCTCTGATGACCCCAACATCTTCAAG	1260
Qy	1525	CACCAGAGAAGTTCAACCCAGACCGTTTCTCTGGATGCAGATGCACGGTTCAGGAAGCAT	1584
Db	1261		
		CACCAGAGAAGTTCAACCCAGACCGTTTCTCTGGATGCAGATGCACGGTTCAGGAAGCAT	1320
Qy	1585	GAGCGTTCCTGCCCTTCTCCTTAGGGAAGCGTCTCTGCCCTTGAGAGGGCTGGCAAAA	1644
Db	1321		
		GAGCGTTCCTGCCCTTCTCCTTAGGGAAGCGTCTCTGCCCTTGAGAGGGCTGGCAAAA	1380
Qy	1645	GGGAGCTCTTCTCTTCTTCCACACCATCCTACAAGGCTTCTCCTTGGAGAGCCCGTGC	1704
Db	1381		
		GGGAGCTCTTCTCTTCTTCCACACCATCCTACAAGGCTTCTCCTTGGAGAGCCCGTGC	1440
Qy	1705	CCGCCGACACCCCTGAGCGCTCAAGCCACCGTCACTGGCGCTTTTCAACATTTCCCCAGCC	1764
Db	1441		
		CCGCCGACACCCCTGAGCGCTCAAGCCACCGTCACTGGCGCTTTTCAACATTTCCCCAGCC	1500
Qy	1765	TTCAGCTGCAAGTCGCTCCCATGACCTTCACTCCACCGCAGACCAGATGAAGGAAG	1824
Db	1501		
		TTCAGCTGCAAGTCGCTCCCATGACCTTCACTCCACCGCAGACCAGATGAAGGAAG	1560
Qy	1825	GCACTTGGAGTGGTGGGTGCCAGACGGTGCCTCAGCCTCAAAGTGGGATGGAC	1884
Db	1561		
		GCACTTGGAGTGGTGGGTGCCAGACGGTGCCTCAGCCTCAAAGTGGGATGGAC	1620
Qy	1885	AGGTTAATGCTCTCACAGTGTACACTGCAGGAGCCACATTTACAGCGCTGCAGTGTGT	1944
Db	1621		
		AGGTTAATGCTCTCACAGTGTACACTGCAGGAGCCACATTTACAGCGCTGCAGTGTGT	1680
Qy	1945	TTCGGAGTCTGTCCACGGCCACACGCTCACTTGACTCATGCTCATAGATGCACAAC	2004
Db	1681		
		TTCGGAGTCTGTCCACGGCCACACGCTCACTTGACTCATGCTCATAGATGCACAAC	1740
Qy	2005	CGCACACCCATACAACTACAAGGCCACAAGCAACTCTCTGGTTAGCTTTCCACAGA	2064
Db	1741		
		CGCACACCCATACAACTACAAGGCCACAAGCAACTCTCTGGTTAGCTTTCCACAGA	1800
Qy	2065	CATAAATATAGTCCATCTGCAATCACAAAGCACATAGCCAGTAAACCACCACTCCCTTG	2124
Db	1801		
		CATAAATATAGTCCATCTGCAATCACAAAGCACATAGCCAGTAAACCACCACTCCCTTG	1860
Qy	2125	GATCTGAGCCCAACGCTGGGAGTCTGGCTGTCACTTCAAAAGCCACAGAAACGGCCAC	2184
Db	1861		
		GATCTGAGCCCAACGCTGGGAGTCTGGCTGTCACTTCAAAAGCCACAGAAACGGCCAC	1920
Qy	2185	ACATGTTACAGCTCACAGCCCTCTCCATTCAITGAACTTCTAGTGTCCCTGTCCCTTG	2244
Db	1921		
		ACATGTTACAGCTCACAGCCCTCTCCATTCAITGAACTTCTAGTGTCCCTGTCCCTTG	1980
Qy	2245	GTGCTGGCACAGGAACAGCATGCCCTCCGGGTCATGCCACCCAGAGACTGTGCGT	2304
Db	1981		
		GTGCTGGCACAGGAACAGCATGCCCTCCGGGTCATGCCACCCAGAGACTGTGCGT	2040
Qy	2305	GTCTATGGCCCCAATCATGCTCCCTCTCTTGGCTACACCACTCTCCCAAGCCTGTGACCA	2364
Db	2041		
		GTCTATGGCCCCAATCATGCTCCCTCTCTTGGCTACACCACTCTCCCAAGCCTGTGACCA	2100



Db	1	CGCGAGCCCTGGAGAGGAGAGAGCGGACCTGGCGAGATGGAGCGGACCGGCACTGG	60	Db	1081	ACCCGCTCCCTTACACGACGCGGTTCTTGCAATGAGGCGACGCGGTGCTGGCGCTGGTG	1140
Qy	325	GGCTGCTGCTGGCGTGGCGTCTCTCTGCTGCTGACGCTGGCGCTGTCCGGACCAGG	384	Qy	1405	CCCATGGGAATACCCGCAACCTCATGCGACCAACCCGCTTCGAGGGGTACACCTCGCC	1464
Db	61	GGCTGCTGCTGGCGTGGCGTGTCTCTGCTGCTGACGCTGGCGTGTCCGGGACCAGG	120	Db	1141	CCCATGGGAATACCCGCAACCTCATGCGACCAACCCGCTTCGAGGGGTACACCTCGCC	1200
Qy	385	GCCCGAGGCACTGCTCCCGCCCGGCGCCACGCGCTACCACTGCTGGGAAACCTCTGTCAG	444	Qy	1465	CAGGGAACGAGGCTTCCCTCCCTCTCTGCTCCATCTCGATGACCCCAACATCTTCAAG	1524
Db	121	GCCCGAGGCACTGCTCCCGCCCGGCGCCACGCGCTACCACTGCTGGGAAACCTCTCTGCA	180	Db	1201	CAGGGAACGAGGCTTCCCTCCCTCTCTGCTCCATCTCGATGACCCCAACATCTTCAAG	1260
Qy	445	CTACGSCCGCGGCGCTGATTCAGGGCTCATGCGCTGAGTAAGAAGTACGACCGGTG	504	Qy	1525	CAACAGAGAGTTCACCCAGACCGTTCTTGATGAGATGAGATGAGATGAGATGAGATGAG	1584
Db	181	CTACGSCCGCGGCGCTGATTCAGGGCTCATGCGCTGAGTAAGAAGTACGACCGGTG	240	Db	1261	CAACAGAGAGTTCACCCAGACCGTTCTTGATGAGATGAGATGAGATGAGATGAGATGAG	1320
Qy	505	TTACCATCTACTGCGGACCTGGCGCTGTGGTCTGTGGTCTGTGGTCTGTGGTCTGTGGT	564	Qy	1585	GAGGCGTTCCTGCTCTCTCTAGGGAACGCTGTCTGCTTGGAGAGCGCTTGGCAAAA	1644
Db	241	TTACCATCTACTGCGGACCTGGCGCTGTGGTCTGTGGTCTGTGGTCTGTGGTCTGTGGT	300	Db	1321	GAGGCGTTCCTGCTCTCTCTAGGGAACGCTGTCTGCTTGGAGAGCGCTTGGCAAAA	1380
Qy	565	CGGAGGCGCTGGAGGTCAGGAGTTCAGCGCGCGGGAACCGTAGCGATGCTG	624	Qy	1645	GCGGAGCTTCTCTCTCTCTCAACCATCTTCAAGCGCTTCTCTCTGAGAGCGCTGCTG	1704
Db	301	CGGAGGCGCTGGAGGTCAGGAGTTCAGCGCGCGGGAACCGTAGCGATGCTG	360	Db	1381	GCGGAGTGTCTCTCTCTCTCAACCATCTTCAAGCGCTTCTCTCTGAGAGCGCTGCTG	1440
Qy	625	GAAAGGACTTTGATGGCCATGGGTTTCTCTCAACGCGGAGCGGTGGAGCGAGCTG	684	Qy	1705	CGCGCGGACACCTCGAGCTCAAGCCACCGTCAAGCGCTTCAACATTCCTCCCGAGCC	1764
Db	361	GAAAGGACTTTGATGGCCATGGGTTTCTCTCAACGCGGAGCGGTGGAGCGAGCTG	420	Db	1441	CGCGCGGACACCTCGAGCTCAAGCCACCGTCAAGCGCTTCAACATTCCTCCCGAGCC	1500
Qy	685	AGGAAGTTTACCATGCTGCTGGGACCTGGGATGGGAGCGGAGGAGCGAGGAG	744	Qy	1765	TTTCCAGCTCAAGTCCGTCCTCACTGACCTTCACTCCACACGACGACGATGAAGAA	1824
Db	421	AGGAAGTTTACCATGCTGCTGGGACCTGGGATGGGAGCGGAGGAGCGAGGAG	480	Db	1501	TTTCCAGCTCAAGTCCGTCCTCACTGACCTTCACTCCACACGACGACGATGAAGAA	1560
Qy	745	CTGATCCAGCGGAGCGCGGTCTGCTGGGAGACATTCAGGGGACAGAAAGACGCCCA	804	Qy	1825	GAACTTGAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG	1884
Db	481	CTGATCCAGCGGAGCGCGGTCTGCTGGGAGACATTCAGGGGACAGAAAGACGCCCA	540	Db	1561	GCAACTTGAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG	1620
Qy	805	TTGCAATCCCTGCTGCTGGGCGGAGCGGCTCAACGTAAGTCTGCTCTCTCTCTCTCT	864	Qy	1885	AGGTTAATGCTTCCAGAGTGTACACTGACGAGCGACCATTTTACACGCTGCGAGTTGT	1944
Db	541	TTGCAATCCCTGCTGCTGGGCGGAGCGGCTCAACGTAAGTCTGCTCTCTCTCTCTCT	600	Db	1621	AGGTTAATGCTTCCAGAGTGTACACTGACGAGCGACCATTTTACACGCTGCGAGTTGT	1680
Qy	865	GGCTCCGCTTCTCTATGAGGATAGGAGTTCAGGCGGCTGCTCCGGGCGAGCTGGT	924	Qy	1945	TTTCCGAGTCTGTCCACCGGCGGACACGCTCACTTGACTCATGCTCTAAGATGCAAC	2004
Db	601	GGCTCCGCTTCTCTATGAGGATAGGAGTTCAGGCGGCTGCTCCGGGCGAGCTGGT	660	Db	1681	TTTCCGAGTCTGTCCACCGGCGGACACGCTCACTTGACTCATGCTCTAAGATGCAAC	1740
Qy	925	ACCTGCTGGAGTCACTCCAGGGGGTTCAGACCTACGAGATGTTCTCTCTCTCTCTCT	984	Qy	2005	CGCACCCCATACAACTTCAAGGGCGCAAAAGCAACTGCTGGGTGTAGCTTCCACAGA	2064
Db	661	ACCTGCTGGAGTCACTCCAGGGGGTTCAGACCTACGAGATGTTCTCTCTCTCTCTCT	720	Db	1741	CGCACCCCATACAACTTCAAGGGCGCAAAAGCAACTGCTGGGTGTAGCTTCCACAGA	1800
Qy	985	CGGCGCTGCGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG	1044	Qy	2065	CATAATATAGTCCATCTGCAATCAAGCACATAGCAGGTAAACCACTCCCTCTG	2124
Db	721	CGGCGCTGCGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG	780	Db	1801	CATAATATAGTCCATCTGCAATCAAGCACATAGCAGGTAAACCACTCCCTCTG	1860
Qy	1045	ACAGTCCGCGAGTGCAGGACACCGGGAACCTGATGCTTCCGGGCGGCGGCGGCGGCG	1104	Qy	2125	GATCTGACGCGCACAGTGGGAGTCTGGTGTCACTTCCACAGCCACAGAAACGGCCAC	2184
Db	781	ACAGTCCGCGAGTGCAGGACACCGGGAACCTGATGCTTCCGGGCGGCGGCGGCGGCG	840	Db	1861	GATCTGACGCGCACAGTGGGAGTCTGGTGTCACTTCCACAGCCACAGAAACGGCCAC	1920
Qy	1105	CTTGTGATGCTCTCTGCTGAAGATGGCAACAGGAGGAAACAAACCCAGGACAGAAATC	1164	Qy	2185	ACATGTTACAGTGCACGCGCTCTCGAATTCATCGAACTTCTCAGTGTCCCTGCTGCTG	2244
Db	841	CTTGTGATGCTCTCTGCTGAAGATGGCAACAGGAGGAAACAAACCCAGGACAGAAATC	900	Db	1921	ACATGTTACAGTGCACGCGCTCTCGAATTCATCGAACTTCTCAGTGTCCCTGCTGCTG	1980
Qy	1165	ACCAACAGAAACATGCTGATGACAGTCAATTTATTTGCTGTTGCTGGGACGATGACGTC	1224	Qy	2245	GTGCTGGCACAGGAAACAGCATGCCCTTCCGGGGTTCATGCCACCCACAGAGACTGCTG	2304
Db	901	ACCAACAGAAACATGCTGATGACAGTCAATTTATTTGCTGTTGCTGGGACGATGACGTC	960	Db	1981	GTGCTGGCACAGGAAACAGCATGCCCTTCCGGGGTTCATGCCACCCACAGAGACTGCTG	2040
Qy	1225	AGCACAGGCTGGCTATACCTTCTGCTCTGATGAATACCTCATGTCCTGCTGCTGCTGCT	1284	Qy	2305	GTCTATGGGCGGCAACTCATGCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2364
Db	961	AGCACAGGCTGGCTATACCTTCTGCTCTGATGAATACCTCATGTCCTGCTGCTGCTGCT	1020	Db	2041	GTCTATGGGCGGCAACTCATGCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2100
Qy	1285	GTACGTGAGAGCTGAATCGGAGCTGGGGCTGGCGGAGGACCAAGACCTTAGGGGACCGT	1344	Qy	2365	CGATGTTACACACACCGGCAACCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2424
Db	1021	GTACGTGAGAGCTGAATCGGAGCTGGGGCTGGCGGAGGACCAAGACCTTAGGGGACCGT	1080	Db	2101	CGATGTTACACACACCGGCAACCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2160
Qy	1345	ACCGGCTCCCTTACACCGAGGCTTCTGATGAGGCGGAGGCTGCTGGGCTGGTG	1404	Qy	2425	GGCTCCCGGAGGATCTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2484
				Db	2161	GGCTCCCGGAGGATCTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2220

```
QY 2485 CACCTCTGCACTGAGCCCTCAGTACCCCTTTTAAAGCACCCTGATTTACCAATGC 2544
Db 2221 CACCTCTGCACTGAGCCCTCAGTACCCCTTTTAAAGCACCCTGATTTACCAATGC 2280
QY 2545 AAACACATCTGGCTGCGATTATGACACAGAGACTTTGGACATACAGAGACCCCTCAGACC 2604
Db 2281 AAACACATCTGGCTGCGATTATGACACAGAGACTTTGGACATACAGAGACCCCTCAGACC 2340
QY 2605 GGAGGAACACTGCGCCAAACCCCAACACAGTGTATGTAACACAGTGGAAAGCGCCCTG 2664
Db 2341 GGAGGAACACTGCGCCAAACCCCAACACAGTGTATGTAACACAGTGGAAAGCGCCCTG 2400
QY 2665 CTGCCCCCTCCACACACATACACACTACACTGATCTACAGCCCTGTTGGCGTCAAGT 2724
Db 2401 CTGCCCCCTCCACACACATACACACTACACTGATCTACAGCCCTGTTGGCGTCAAGT 2460
QY 2725 CCCCACTAGACCCAGTGGAGGGGTAGAGACCAAGTAGGGCCAGTTTCCAAATTCACCC 2784
Db 2461 CCCCACTAGACCCAGTGGAGGGGTAGAGACCAAGTAGGGCCAGTTTCCAAATTCACCC 2520
QY 2785 TGTGAGGAGTACCGGAGTCTGAGCTTCTGTTGACCTTAAGGTCGGCTTGGGAATTA 2844
Db 2521 TGTGAGGAGTACCGGAGTCTGAGCTTCTGTTGACCTTAAGGTCGGCTTGGGAATTA 2580
QY 2845 AAGTTTCTTCTGGCCCTTAGCCT 2868
Db 2581 AAGTTTCTTCTGGCCCTTAGCCT 2604

RESULT 7
AX876954
LOCUS AX876954 2513 bp DNA linear PAT 17-DEC-2003
DEFINITION Sequence 11859 from Patent EP1074617.
ACCESSION AX876954
VERSION AX876954.1 GI:40031690
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Ota, T., Isogai, T., Nishikawa, T., Hayashi, K., Saito, K., Yamamoto, J.,
Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K. and Otsuki, T.
TITLE Primers for synthesising full-length cDNA and their use
JOURNAL Patent: EP 1074617-A 11859 07-FEB-2001;
FEATURES
source
1. .2513
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
CDS
56..1750
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAB89506.1"
/db_xref="GI:40031691"
/translation="MEATGFWALLLALLLLTLALSLGTRARGHLPPGPTPLPLGN
LIQLRGALYGLMLRSKKYGVFTIYLGFWPVLVVGQEAAREALGQAEFRSGR
TVAMLEGTDGHGVFFSNGERWQLKFTMLALDLGMKREGEELIQAEARCLVTF
QGTGEPFPDPLSLAQAQSNVCSLLFLGRFSYEDKEFOAVVRAAGTLLGVSSQSGQ
TYEMFNGFLPLPGPHQKLIHVVSTIAFTVROVQHQNLDAAGPARDLVAPFLKM
AOEONQTEFTKNMLTVIYLLFAGTMTVSTVGYTLMLLMKYIPIHQKRVREELN
ELGAGAPSLGDRTRIPYDAVLHEAQRLLVPMGIPRTLMRTTFRGYTLPPQSTEV
FLILGSLIHDNFIKHPDEFNPDRLDADGRFKHEAFPLFSIGKRVCLGEGIAKEL
FLFPTTILQAFSLSEPCDITSLKPTVSGLENIIPASTVGMDRVNVSVYTAGSHY
TPAVFRSLSHGPHALHTAAKXHNRTPIHNVKHKATAGLAFHRHKYSPSALTST"

Query Match 81.5%; Score 2399.4; DB 6; Length 2513;
Best Local Similarity 96.1%; Pred. No. 0;
Matches 2512; Conservative 0; Mismatches 1; Indels 102; Gaps 1;
```

1081	Db	 AAGCTAGGGGACCGTATACCGCGCTCCCTTACACCGACGGGTTCTGCATAGGCGCAGCG	1140
1389	Qy	GCTGCTGGCGCTGGTGGCCATGGGAATACCCCGCACCTCATGGGACACACCGCTTCG	1448
1141	Db	GCTGCTGGCGCTGGTGGCCATGGGAATACCCCGCACCTCATGGGACACACCGCTTCG	1200
1449	Qy	AGGATACACCTGCCCCAGGGCACGGAGTCTTCCCCCTCTCTTGGTTCATCTCGCATGA	1508
1201	Db	AGGGTACACCTGCCCCAGGGCACGGAGTCTTCCCCCTCTCTTGGTTCATCTCGCATGA	1260
1509	Qy	CCCCAACATCTTCAAGCACCCAGAGAAGTTCAACCCAGACCGTTTCTTGGATCAGATGG	1568
1261	Db	CCCCAACATCTTCAAGCACCCAGAGAAGTTCAACCCAGACCGTTTCTTGGATCAGATGG	1320
1569	Qy	ACGGTTCAGGAAGCATGAGGGTTCTCGCCCTCTCTTACCAACCATCTCAAGGCTTCTC	1628
1321	Db	ACGGTTCAGGAAGCATGAGGGTTCTCGCCCTCTCTTACCAACCATCTCAAGGCTTCTC	1380
1629	Qy	AGAGGGCTTGGCAAAAGCGGAGTCTTCTCTTCTTCCACCAACCATCTCAAGGCTTCTC	1688
1381	Db	AGAGGGCTTGGCAAAAGCGGAGTCTTCTCTTCTTCCACCAACCATCTCAAGGCTTCTC	1440
1689	Qy	CCTGGAGAGCCCTGCCCCGGGACACCTTGAGCCTCAAGCCACCGTCAAGTGGCCTTTT	1748
1441	Db	CCTGGAGAGCCCTGCCCCGGGACACCTTGAGCCTCAAGCCACCGTCAAGTGGCCTTTT	1500
1749	Qy	CAACATTTCCCCAGCCTTCCAGCTGCGAAGTCCGTCCTCCACTGACCTTCACACCAGCA	1808
1501	Db	CAACATTTCCC-----	1510
1809	Qy	GACCAGATGAAGGAAGGCAACTTGGAAATGGTGGTGCCCGAGGACGGTGCTCCAGCCTC	1868
1511	Db	-----CGAGCTC	1518
1869	Qy	AACAGTGGGCATGACAGGGTTAATGTCTCCAGAGTGATACACTGCAGGACGCCACATTTA	1928
1519	Db	AACAGTGGGCATGACAGGGTTAATGTCTCCAGAGTGATACACTGCAGGACGCCACATTTA	1578
1929	Qy	CACGCTGCAGTTGTTTTCGGAGTCTGTCCACGGGCCACACGCTCACTTGACTCATGC	1988
1579	Db	CACGCTGCAGTTGTTTTCGGAGTCTGTCCACGGGCCACACGCTCACTTGACTCATGC	1638
1989	Qy	TGCTAAGATGCACAACCGCACACCCATACACAACACTCAAGGGCCACAAAGCAACTGCTGG	2048
1639	Db	TGCTAAGATGCACAACCGCACACCCATACACAACACTCAAGGGCCACAAAGCAACTGCTGG	1698
2049	Qy	GTTAGCTTTCCACAGACATAAATATAGTCCATCTGCAATCACAAGCACATAGCCAGTAA	2108
1699	Db	GTTAGCTTTCCACAGACATAAATATAGTCCATCTGCAATCACAAGCACATAGCCAGTAA	1758
2109	Qy	CCCACCAACTCCCTGGATCTGACGCCACACAGTGGGAGTCTGGCTGTCACTTCAAG	2168
1759	Db	CCCACCAACTCCCTGGATCTGACGCCACACAGTGGGAGTCTGGCTGTCACTTCAAG	1818
2169	Qy	CCACAGAAACGGCCACACATGTTTCACAGCTCACACGCCCTCTCCATTTCGAACTTCTC	2228
1819	Db	CCACAGAAACGGCCACACATGTTTCACAGCTCACACGCCCTCTCCATTTCGAACTTCTC	1878
2229	Qy	AGTGTCCCTGTCCCTGGTGGCCAGGGAAACAGCATGCCCTCCGGGTCAAGCCA	2288
1879	Db	AGTGTCCCTGTCCCTGGTGGCCAGGGAAACAGCATGCCCTCCGGGTCAAGCCA	1938
2289	Qy	CCCAGAGACTGTGGTGTCTATGGCCCAACTCATGTCTCCTCTCTTGGCTACACCACTC	2348
1939	Db	CCCAGAGACTGTGGTGTCTATGGCCCAACTCATGTCTCCTCTCTTGGCTACACCACTC	1998
2349	Qy	TCCCAGCCTGTGACACACCGATGTCCACACACCCCAACCACTGTCTCCACACAGCTACCCA	2408
1999	Db	TCCCAGCCTGTGACACACCGATGTCCACACACCCCAACCACTGTCTCCACACAGCTACCCA	2058
2409	Qy	CGTACGACATGCTCTGGCTCCCCAGAGTATCTTTCCACTGAGACACGCCGCCCCACAG	2468

Db	2059	CGTACGACATCGTCCCGGCTCCCGAGAGTATCTTCCCACTGAGACACGCCGCCCCACAG	2111
Qy	2469	AGGCACAGTCCCGACGACCACTCTGCACTGAGCGCCCTCAGTCACCCCTTTTAAAGCACCC	2528
Db	2119	AGGCACAGTCCCGACGACCACTCTGCACTGAGCGCCCTCAGTCACCCCTTTTAAAGCACCC	2178
Qy	2529	TGATTCTACAAATGCAAAACATCTCTGGGTCTGGGATTATGCAACAGAGACTTTGGACATA	2588
Db	2179	TGATTCTACAAATGCAAAACATCTCTGGGTCTGGGATTATGCAACAGAGACTTTGGACATA	2238
Qy	2589	CGAGGACCTTCAGACCGGAGGACACCTGCGCCACCCCAACAGCTGCTTATGTAACCAACG	2648
Db	2239	CGAGGACCTTCAGACCGGAGGACACCTGCGCCACCCCAACAGCTGCTTATGTAACCAACG	2298
Qy	2649	TGAAAAGCGGCCCTGCTGCGCCCTCCACACACACATACACTCAGTATCTACAGCCCC	2708
Db	2299	TGAAAAGCGGCCCTGCTGCGCCCTCCACACACACATACACTCAGTATCTACAGCCCC	2358
Qy	2709	TGTTGCGGCTCAGAGTCCCACTAGACCCAGTGAAGGGGTTAGAGACCAAGTAGGGGCC	2768
Db	2359	TGTTGCGGCTCAGAGTCCCACTAGACCCAGTGAAGGGGTTAGAGACCAAGTAGGGGCC	2418
Qy	2769	AGTTTCCAATTCACCCCTGTGAGGAGTGAGCGGATCTGACGTTTCCCTTGACTTAAGGG	2828
Db	2419	AGTTTCCAATTCACCCCTGTGAGGAGTGAGCGGATCTGACGTTTCCCTTGACTTAAGGG	2478
Qy	2829	TCCGGCTTGGAATTAAGTTTGTCTGGCCTTT	2863
Db	2479	TCCGGCTTGGAATTAAGTTTGTCTGGCCTTT	2513
RESULT 8			
BD156406		2513 bp	linear
LOCUS	BD156406	2513 bp	DNA
DEFINITION	Primer for synthesizing full-length cDNA and use thereof.		
ACCESSION	BD156406.1		
VERSION	BD156406.1 GI:27862164		
KEYWORDS	JP 2002191363-A/11249.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 2513)		
TITLE	Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J., Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.		
JOURNAL	Primer for synthesizing full-length cDNA and use thereof		
COMMENT	Patent: JP 2002191363-A 11249 09-JUL-2002; HELIX RESEARCH INSTITUTE OS Homo sapiens (human) PN JP 2002191363-A/11249 PD 09-JUL-2002 PF 28-JUL-2000 JP 2000280990 PI TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU SAITO, PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU, PI KEIICHI NAGAI, TETSUJI OTSUKI PC C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/PC 10, PC C12P21/02,C12Q1/68/C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC Primer for synthesizing full-length cDNA and use thereof FH Key		
FEATURES	FT CDS	Location/Qualifiers	
source	1..2513	Location/Qualifiers	
ORIGIN	/organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606"		
Query Match	81.5%;	Score 2399.4;	DB 6; Length 2513;
Best Local Similarity	96.1%;	Prod. No. 0;	
Matches 2512;	Conservative 0;	Mismatches 1;	Indels 102; Gaps 1







QY 969 GTTCTCTGGTTCCTGCGGCCCTGCCAGGCCCCACAAAGCAGCTCCTCCACCAAGCTCAG 1028  
Db 721 GTTCTCTGGTTCCTGCGGCCCTGCCAGGCCCCACAAAGCAGCTCCTCCACCAAGCTCAG 780  
QY 1029 CACCTTGGCTGCTTCAAGTCCGGCAGGTGAGCAGCACCAGGGGAACTGGATGCTTC 1088  
Db 781 CACCTTGGCTGCTTCAAGTCCGGCAGGTGAGCAGCACCAGGGGAACTGGATGCTTC 840  
QY 1089 GGGCCCCCAGCTGACCTTGTGATGCTTCTGCTGAAGTGGCAGAGGAGAAACAAA 1148  
Db 841 GGGCCCCCAGCTGACCTTGTGATGCTTCTGCTGAAGTGGCAGAGGAGAAACAAA 900  
QY 1149 CCCAGGCACAGAAATTCACCAACAAGAACATGCTGATGACAGTCAATTTATTTGCTGTTTC 1208  
Db 901 CCCAGGCACAGAAATTCACCAACAAGAACATGCTGATGACAGTCAATTTATTTGCTGTTTC 960  
QY 1209 TGGGACGATGACGCTGACGACACACAGGTGCGGTATACCCCTCTGCTGATGAATACCC 1268  
Db 961 TGGGACGATGACGCTGACGACACACAGGTGCGGTATACCCCTCTGCTGATGAATACCC 1020  
QY 1269 TCATGTCCTCAAGTGGGTACGCTGAGGAGCTGAATCGGAGCTGGGGGTGSCCAGCAC 1328  
Db 1021 TCATGTCCTCAAGTGGGTACGCTGAGGAGCTGAATCGGAGCTGGGGGTGSCCAGCAC 1080  
QY 1329 AAGCTTAGGGGACCGGTACCCGCCCTCCCTTACACCGACGCGGTCTGCAATGAGGCGCAGC 1388  
Db 1081 AAGCTTAGGGGACCGGTACCCGCCCTCCCTTACACCGACGCGGTCTGCAATGAGGCGCAGC 1140  
QY 1389 GCTGCTGGCGTGTGCGCATGCGGAATACCCCGACCCCTCATGCGGACACCCGCTTCGG 1448  
Db 1141 GCTGCTGGCGTGTGCGCATGCGGAATACCCCGACCCCTCATGCGGACACCCGCTTCGG 1200  
QY 1449 AGGATCACCTGCCCGCAGGACGAGGCTTCCCTCTCTGCTGCTGCTGCTGCTGCTGCTG 1508  
Db 1201 AGGATCACCTGCCCGCAGGACGAGGCTTCCCTCTCTGCTGCTGCTGCTGCTGCTGCTG 1260  
QY 1509 CCCCAACATCTTCAAGCACCCAGAGAGTTCAACCCAGACCGTTTCTGGAATGAGATGG 1568  
Db 1261 CCCCAACATCTTCAAGCACCCAGAGAGTTCAACCCAGACCGTTTCTGGAATGAGATGG 1320  
QY 1569 ACGTTTCAAGAGATGAGGCTTCTGCTGCTTCTTCAAGGAGGCTGCTGCTGCTGCTGCTG 1628  
Db 1321 ACGTTTCAAGAGATGAGGCTTCTGCTGCTTCTTCAAGGAGGCTGCTGCTGCTGCTGCTG 1380  
QY 1629 AGAGGGCTGCAAAAGGGAGCTTCTTCTTCTTCAACACCATCTCAAGCCTTCTC 1688  
Db 1381 AGAGGGCTGCAAAAGGGAGCTTCTTCTTCTTCAACACCATCTCAAGCCTTCTC 1440  
QY 1689 CTTGGAGAGCCGTCGCCCGCGGACACCTGAGCTCAAGCCCAACCTCAGTGGCCTTTT 1748  
Db 1441 CTTGGAGAGCCGTCGCCCGCGGACACCTGAGCTCAAGCCCAACCTCAGTGGCCTTTT 1500  
QY 1749 CAACATTCCTCCAGCTTCCAGCTGCAAGTCCGTCCTCACTGACCTTCACTCCACCAACGCA 1808  
Db 1501 CAACATTCCTCCAGCTTCCAGCTGCAAGTCCGTCCTCACTGACCTTCACTCCACCAACGCA 1510  
QY 1809 GACCAGATGAAGGAGCACTTGGAGTGGTGGTCCCGAGGAGCTGCTCCAGCCTC 1868  
Db 1511 GACCAGATGAAGGAGCACTTGGAGTGGTGGTCCCGAGGAGCTGCTCCAGCCTC 1518  
QY 1869 AACAGTGGGATGAGCAGAGGTTAATGCTCCAGAGTGTACACTGACGAGGAGGACATTTA 1928  
Db 1519 AACAGTGGGATGAGCAGAGGTTAATGCTCCAGAGTGTACACTGACGAGGAGGACATTTA 1578  
QY 1929 CAGCTTGCAGTGTGTTTCCGGAGTGTGCTCCAGGCGCCACAGCTCACTTGACTCATGC 1988  
Db 1579 CAGCTTGCAGTGTGTTTCCGGAGTGTGCTCCAGGCGCCACAGCTCACTTGACTCATGC 1638  
QY 1989 TGCTAGATGCAACCGCACCCATACACAACTACAGGGGACAAAGCACTGCTGG 2048  
Db 1639 TGCTAGATGCAACCGCACCCATACACAACTACAGGGGACAAAGCACTGCTGG 1698  
QY 2049 GTTAGCTTTCCACAGACATAAATATAGTCCATCTGCAATCACAAGCACATAGCCAGGTAA 2108

Db 1699 GTTAGCTTTCCACAGACATAAATATAGTCCATCTGCAATCACAAGCACATAGCCAGGTAA 1758  
QY 2109 CCCACAACTCCCTGGATCTGCAGCCCAACAGTGGGAGTCTGGCTGTACCTTCCACAG 2168  
Db 1759 CCCACAACTCCCTGGATCTGCAGCCCAACAGTGGGAGTCTGGCTGTACCTTCCACAG 1818  
QY 2169 CCACAGAAACGGGACACATGTTTCCAGCTCAGCGCCCTTCCATTCATCGAACTTCTC 2228  
Db 1819 CCACAGAAACGGGACACATGTTTCCAGCTCAGCGCCCTTCCATTCATCGAACTTCTC 1878  
QY 2229 AGTGTCCCTGTCCTGGTGGCTGACAGGAAAGAGTGGCCCTCCGGGGTCAAGCA 2288  
Db 1879 AGTGTCCCTGTCCTGGTGGCTGACAGGAAAGAGTGGCCCTCCGGGGTCAAGCA 1938  
QY 2289 CCCAGAGACTGTGCTGTCTATGCGCCCAACTCATGCTCCCTCTCTTGGCTACACACTC 2348  
Db 1939 CCCAGAGACTGTGCTGTCTATGCGCCCAACTCATGCTCCCTCTCTTGGCTACACACTC 1998  
QY 2349 TCCAGAGCTGTGACACACCGATGTCCACACACCCCAACCACTTGTCCACACAGCTTACCCA 2408  
Db 1999 TCCAGAGCTGTGACACACCGATGTCCACACACCCCAACCACTTGTCCACACAGCTTACCCA 2058  
QY 2409 CGTAGCAGATCGTCTGCTGCTCCCGAGAGTATCTTCCCACTGAGACACGCGCCGCCACAG 2468  
Db 2059 CGTAGCAGATCGTCTGCTGCTCCCGAGAGTATCTTCCCACTGAGACACGCGCCGCCACAG 2118  
QY 2469 AGGCACAGTCTCCCAACGACCTCTGCAACTGAGCGCTCAGTCAACCCCTTTTAAAGCACCC 2528  
Db 2119 AGGCACAGTCTCCCAACGACCTCTGCAACTGAGCGCTCAGTCAACCCCTTTTAAAGCACCC 2178  
QY 2529 TGATTTACCAATGCAAAACACATCTGGTCTGCGATTATGCAAGAGACTTTGACACATA 2588  
Db 2179 TGATTTACCAATGCAAAACACATCTGGTCTGCGATTATGCAAGAGACTTTGACACATA 2238  
QY 2589 CGAGGACCTCTAGACCGGAGGACACCTGCCCAACACCTGCTGCTTATGTAACACAG 2648  
Db 2239 CGAGGACCTCTAGACCGGAGGACACCTGCCCAACACCTGCTGCTTATGTAACACAG 2298  
QY 2649 TGGAAAGCGGCCCTCTGCTGCCCTTCCACACACATACACTCACTGATCTACAGCCCC 2708  
Db 2299 TGGAAAGCGGCCCTCTGCTGCCCTTCCACACACATACACTCACTGATCTACAGCCCC 2358  
QY 2709 TGTTGGCGTCAAGTCTCCCACTAGACCCAGTGGAGGGTTAGAGACCAAGTAGGGGCC 2768  
Db 2359 TGTTGGCGTCAAGTCTCCCACTAGACCCAGTGGAGGGTTAGAGACCAAGTAGGGGCC 2418  
QY 2769 AGTTTCCAATTCACCTGTCAGGAGTGAGCGGATCTGAGCTTCTTGTGACTTAAGG 2828  
Db 2419 AGTTTCCAATTCACCTGTCAGGAGTGAGCGGATCTGAGCTTCTTGTGACTTAAGG 2478  
QY 2829 TCCGGCTTGGGAATTAAGTTTCTGCGCTTT 2863  
Db 2479 TCCGGCTTGGGAATTAAGTTTCTGCGCTTT 2513

## RESULT 10

AX780433

LOCUS

DEFINITION

AX780433

ACCESSION

AX780433.1

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

AX780433 2261 bp DNA linear PAT 14-JUL-2003  
Sequence 2590 from Patent WO03039443.

AX780433.1 GI:32697427

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1

Haerlath, T., Schoch, C., Kern, W., Kohlmann, A., Schnittger, S.,

Dugas, M., Bils, R., Brors, B. and Mergenthaler, S.

Novel genetic markers for leukemias

Patent: WO 03039443-A 2590 15-MAY-2003;

Deutsches Krebsforschungszentrum (DKFZ);







Db	1081	CAGCGCTGCTGGCGTGTGGCCATGGGAATACCCGACACCTCATGCGGACACCCGC	1140
Qy	1444	TTCCGAGGATACACCTGCCCCAGGGACGAGGTCTTCCCTCTCTTGGCTCCATCTTG	1503
Db	2141	TTCCGAGGATACACCTGCCCCAGGGACGAGGTCTTCCCTCTCTTGGCTCCATCTTG	1200
Qy	1504	CATGACCCCAACATCTTCAAGACCCAGAGATTCAACCCAGACCGTTTCTGGATGCA	1563
Db	1201	CATGACCCCAACATCTTCAAGACCCAGAGATTCAACCCAGACCGTTTCTGGATGCA	1260
Qy	1564	GATGACCGTTTCAAGACCATGAGGCTTCTGCTCTTCTTAGGGAAGGTGCTGTC	1623
Db	1261	GATGACCGTTTCAAGACCATGAGGCTTCTGCTCTTCTTAGGGAAGGTGCTGTC	1320
Qy	1624	CTTGAGAGGGCTGGCAAAAGCGAGCTTCTCTCTTCTTCAACACCATCTTACAAGCC	1683
Db	1321	CTTGAGAGGGCTGGCAAAAGCGAGCTTCTCTCTTCTTCAACACCATCTTACAAGCC	1380
Qy	1684	TTCTCCTGGAGAGCGGTGCGCGGACACCTGAGCTCAAGCCACCGTCACTGGC	1743
Db	1381	TTCTCCTGGAGAGCGGTGCGCGGACACCTGAGCTCAAGCCACCGTCACTGGC	1440
Qy	1744	CTTTTCAACATTTCCCGAGCTTCCAGCTGCAAGTCCGTCCCACTGACCTTCACTCCAC	1803
Db	1441	CTTTTCAACATTTCCCGAGCTTCCAGCTGCAAGTCCGTCCCACTGACCTTCACTCCAC	1500
Qy	1804	ACGACACCATGTAAGGAAGCACTTGGAAAGTGTGGTGGCCAGGCGGTCTTCA	1863
Db	1501	ACGACACCATGTAAGGAAGCACTTGGAAAGTGTGGTGGCCAGGCGGTCTTCA	1560
Qy	1864	GCCTCAACAGTGGGATGACAGGTTAATGTCTCCAGAGTGTACACTCAGGACGCCAC	1923
Db	1561	GCCTCAACAGTGGGATGACAGGTTAATGTCTCCAGAGTGTACACTCAGGACGCCAC	1620
Qy	1924	ATTTACACGCTCGAGTGTGTTTCCGAGTCTGTCCACGCGCCACACGCTCACTTGACT	1983
Db	1621	ATTTACACGCTCGAGTGTGTTTCCGAGTCTGTCCACGCGCCACACGCTCACTTGACT	1680
Qy	1984	CATGCTGTAAGATGACAAACCGCACACCATACAACTTACAAAGGCGCCAAAGCACT	2043
Db	1681	CATGCTGTAAGATGACAAACCGCACACCATACAACTTACAAAGGCGCCAAAGCACT	1740
Qy	2044	GCTGGGTAGCTTTCACAGACATAATAGTTCATCTGCAATCAAGGCGCCATAGCCA	2103
Db	1741	GCTGGGTAGCTTTCACAGACATAATAGTTCATCTGCAATCAAGGCGCCATAGCCA	1800
Qy	2104	GGTAACCCCAACTCCCTGGATCTGACGCCACACAGTGGGAGTCTGGCTGTCACTTC	2163
Db	1801	GGTAACCCCAACTCCCTGGATCTGACGCCACACAGTGGGAGTCTGGCTGTCACTTC	1860
Qy	2164	ACAAGCCACAGAAACGGGCGACATGTTTCAAGCTTCAAGCGCTTCCATTCATGAAAC	2223
Db	1861	ACAAGCCACAGAAACGGGCGACATGTTTCAAGCTTCAAGCGCTTCCATTCATGAAAC	1920
Qy	2224	TTCTCAGTGTCCCTGTCCCTGGTGGCCACAGGGAAGAGTGCCTCCGCGGTCA	2283
Db	1921	TTCTCAGTGTCCCTGTCCCTGGTGGCCACAGGGAAGAGTGCCTCCGCGGTCA	1980
Qy	2284	TGCCACCCAGAGACTGTGCTGTCTATGGCCCCAACTCATGCTCCCTCTCTTGGCTACAC	2343
Db	1981	TGCCACCCAGAGACTGTGCTGTCTATGGCCCCAACTCATGCTCCCTCTCTTGGCTACAC	2040
Qy	2344	CATCTCCAGCTGTGACACCATGTTCCACACACCCCAACCACTGTGTCACAGCT	2403
Db	2041	CATCTCCAGCTGTGACACCATGTTCCACACACCCCAACCACTGTGTCACAGCT	2100
Qy	2404	ACCCAGTACAGACTGTCTGGTCTCCCGAGAGTATCTTCCCACTGAGACACGCGGCC	2463
Db	2101	ACCCAGTACAGACTGTCTGGTCTCCCGAGAGTATCTTCCCACTGAGACACGCGGCC	2160
Qy	2464	CACAGAGGCAAGTCCCAGCCACTCTGCAACTGACGCTCAGTCAACCCCTTTTAAAG	2523
Db	2161	CACAGAGGCAAGTCCCAGCCACTCTGCAACTGACGCTCAGTCAACCCCTTTTAAAG	2220
Qy	2524	CACCCCTGA	2531
Db	2221	CACCATGA	2228
RESULT 12			
AX552231			
LOCUS	AX552231	2114 bp	DNA linear PAT 27-NOV-2002
DEFINITION	Sequence 140 from Patent WO0162927.		
ACCESSION	AX552231		
VERSION	AX552231.1	GI:25896469	
KEYWORDS	Homo sapiens (human)		
SOURCE	Homo sapiens		
ORGANISM	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
REFERENCE	1		
AUTHORS	Banville, S.C., Greenawalt, L.B., Lincoln, S.E., Stockdreher, T.K., Amshey, S., Chang, S.C., Chen, W., D'Sa, S.A., Dam, T.C., Liu, T.F., Rosen, B.H., Russo, F.D., Spiro, P.A., Bradley, D.L., Chen, A., Cohen, H.J., Dafo, A., Daniels, S.E., Dufour, G.E., Flores, V., Fong, W.T., Hodgson, D.M., Jackson, S., Jones, A.L., Panzer, S., Roseberry, A.M., Shah, P., Wright, R.J., Yap, P.E., Yu, J.Y., Bratcher, S.R., Chalup, M.S., Dahl, C.R. and Hillman, J.L.		
TITLE	Polypeptides and corresponding polynucleotides for diagnostics and therapeutics		
JOURNAL	Patent: WO 0162927-A 140 30-AUG-2001;		
FEATURES	Incyte Genomics, Inc. (US)		
source	1. .2114		
	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:9606"		
	/note="Incyte ID No: LI:337818.1:2000FEB01"		
ORIGIN			
Query Match	58.3%	Score 1717.4;	DB 6; Length 2114;
Best Local Similarity	99.3%	Pred. No. 4.9e-305;	
Matches 1735;	Conservative 0;	Mismatches 11;	Indels 1; Gaps 1;
Qy	1123	CTGAAGATGGCAGGAGGAAACAAACCCAGGACAGAAATTCACCAAGAAACATGCTG	1182
Db	368	CTTTTATGATGATGAGGAAACAAACCCAGGACAGAAATTCACCAAGAAACATGCTG	427
Qy	1183	ATGACAGTCAATTTATTTGCTTTGCTGGAGATCAGCGTACAGCCAGGTCGCTAT	1242
Db	428	ATGACAGTCAATTTATTTGCTTTGCTGGAGATCAGCGTACAGCCAGGTCGCTAT	487
Qy	1243	ACCTCTCTGCTCTGATGAAATACCTCTCATGTCCAAAAGTGGGTACGTGAGGAGCTGAAT	1302
Db	488	ACCTCTCTGCTCTGATGAAATACCTCTCATGTCCAAAAGTGGGTACGTGAGGAGCTGAAT	547
Qy	1303	CGGAGCTGGGGCTGGCCAGGACCAAGCTTAGGGACCGTACCGCTCTCCCTTACACC	1362
Db	548	CGGAGCTGGGGCTGGCCAGGACCAAGCTTAGGGACCGTACCGCTCTCCCTTACACC	607
Qy	1363	GACGCGGTTCTGATGAGGCGACGCGTCTGCGCTGTGTCCTATGGGAATACCCCGC	1422
Db	608	GACGCGGTTCTGATGAGGCGACGCGTCTGCGCTGTGTCCTATGGGAATACCCCGC	667
Qy	1423	ACCTCAT--GCGGACCAACCGCTTCCGAGGGTACACCTGCCCCAGGACCGAGGTCTT	1481
Db	668	ACCTCATGCGGACCAACCGCTTCCGAGGGTACACCTGCCCCAGGACCGAGGTCTT	727
Qy	1482	CCCCTCTTGGGTCTCATCTCTGATGACCCCAACATCTTCAAGCACCCAGAGGTTCAA	1541
Db	728	CCCCTCTTGGGTCTCATCTCTGATGACCCCAACATCTTCAAGCACCCAGAGGTTCAA	787
Qy	1542	CCGAGACCGTTTCTTGATGACGAGTTCAGGAAGCATGAGCGGTTCTTCCGCTT	1601
Db	788	CCGAGACCGTTTCTTGATGACGAGTTCAGGAAGCATGAGCGGTTCTTCCGCTT	847



Db 361 GGGAGCGGTGAGGACGTGAGGAATTTACATGCTTGCTGCGGACCTGGCATG 420  
QY 724 GGAAGCGAAGAGCGAGAGCTGATCAGGGCGAGGCCGGTGTCTGTGTGAGACATTC 783  
Db 421 CGGAAGCGAAGAGCGAGAGCTGATCAGGGCGAGGCCGGTGTCTGTGTGAGACATTC 480  
QY 784 CAGGGACAGAGAGGAGCGCCATTGATCCCTCCCTGCTGCGCCAGCCACCTCCAAC 843  
Db 481 CAGGGACAGAGAGGAGCGCCATTGATCCCTCCCTGCTGCGCCAGCCACCTCCAAC 540  
QY 844 GTAGTCTGCTCCCTCTCTTTGGCCCTCCGCTTCTCTATGAGGATAAGGAGTTCAGGCC 903  
Db 541 GTAGTCTGCTCCCTCTCTTTGGCCCTCCGCTTCTCTATGAGGATAAGGAGTTCAGGCC 600  
QY 904 GTGGTCCGGGAGCTGGTGGTACCTCTGCTGGAGTCAGTCCAGGGGGGTGAGACCTAC 963  
Db 601 GTGGTCCGGGAGCTGGTGGTACCTCTGCTGGAGTCAGTCCAGGGGGGTGAGACCTAC 660  
QY 964 GAGATGTTCTCTGCTGCTGCTGCGGCCCTGCCAGGCCCCCAACAGCAGCTCTCCACCCAC 1023  
Db 661 GAGATGTTCTCTGCTGCTGCTGCGGCCCTGCCAGGCCCCCAACAGCAGCTCTCCACCCAC 720  
QY 1024 GTACAGCACCCTGCTGCTGCTTACAGTCCGGCAGGTGCAGCAGCACAGGGGAACCTGGAT 1083  
Db 721 GTACAGCACCCTGCTGCTGCTTACAGTCCGGCAGGTGCAGCAGCACAGGGGAACCTGGAT 780  
QY 1084 GCTTCGGGCCCCGACGTGACCTTGTGATGCTTCTGCTGAGAGTGCACAGGAGAA 1143  
Db 781 GCTTCGGGCCCCGACGTGACCTTGTGATGCTTCTGCTGAGAGTGCACAGGAGAA 840  
QY 1144 CAAACCCAGGACAGAAATTCACCAACAGAAATTCACCAACAGAAATTCATGATGCTG 1203  
Db 841 CAAACCCAGGACAGAAATTCACCAACAGAAATTCACCAACAGAAATTCATGATGCTG 900  
QY 1204 TTTGCTGGGACGATGCGGTGACGACCAACCGGTGCTATACCTCTCTGCTCTGATGAA 1263  
Db 901 TTTGCTGGGACGATGCGGTGACGACCAACCGGTGCTATACCTCTCTGCTCTGATGAA 960  
QY 1264 TACCCTCATGTCCAAAGTGGTACGTGAGGAGTGAATCGGAGCTGGGGCTGGCCAG 1323  
Db 961 TACCCTCATGTCCAAAGTGGTACGTGAGGAGTGAATCGGAGCTGGGGCTGGCCAG 1020  
QY 1324 GCACCAAGCCTAGGGACCGTACCGCCTCCCTTACACGACCGCGTTCGTGATGAGGG 1383  
Db 1021 GCACCAAGCCTAGGGACCGTACCGCCTCCCTTACACGACCGCGTTCGTGATGAGGG 1080  
QY 1384 CAGCGCTGCTGCGCTGGTGCCTATGGGAATACCGCGACCTCATGCGGACCAACCCGC 1443  
Db 1081 CAGCGCTGCTGCGCTGGTGCCTATGGGAATACCGCGACCTCATGCGGACCAACCCGC 1140  
QY 1444 TTTCCGAGGTACACCTGCCCCAGGGCAGGAGTCTTCCCTCTCTGCTCCATCCTG 1503  
Db 1141 TTTCCGAGGTACACCTGCCCCAGGGCAGGAGTCTTCCCTCTCTGCTCCATCCTG 1200  
QY 1504 CATGACCCCAACATCTTCAAGACACCCAGAGAGTTCACACCGACCGCTTTCTGGATGCA 1563  
Db 1201 CATGACCCCAACATCTTCAAGACACCCAGAGAGTTCACACCGACCGCTTTCTGGATGCA 1260  
QY 1564 GATGACCGTTCAGGAAGCATGAGGGCTTCTGCGCTTCTCTTTAGGGAAGCGTGTGTC 1623  
Db 1261 GATGACCGTTCAGGAAGCATGAGGGCTTCTGCGCTTCTCTTTAGGGAAGCGTGTGTC 1320  
QY 1624 CTTTGGAGGGCTGGCAAGAGCGAGCTCTTCTCTTCTTCAACCACTCTTCAAGCC 1683  
Db 1321 CTTTGGAGGGCTGGCAAGAGCGAGCTCTTCTCTTCTTCAACCACTCTTCAAGCC 1380  
QY 1684 TTTCTCCCTGGAGAGCCCGTCCCGCGGACACCTTGAGCCTCAAGCCACCGTCAGTGGC 1743  
Db 1381 TTTCTCCCTGGAGAGCCCGTCCCGCGGACACCTTGAGCCTCAAGCCACCGTCAGTGGC 1440  
QY 1744 CTTTCAACATTTCCCGAGCCTTCAGTGCAGTCCGTCCTCCACTGACCTTCACTCCAC 1803

Db 1441 CTTTCAACATTTCCCGAGCCTTCCAGCTTCCAGTGCAGTCCGTCACCTGACCTTCACTCCACC 1500  
QY 1804 ACGCAGACCAGATGA 1818  
Db 1501 ACGCAGACCAGATGA 1515  
RESULT 14  
AC011510 129402 bp DNA linear PRI 23-OCT-2000  
LOCUS Homo sapiens chromosome 19 clone CTD-2195B23, complete sequence.  
AC011510  
AC011510.7 GI:10947023  
VERSION HTG.  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
1 (bases 1 to 129402)  
DOE Joint Genome Institute and Stanford Human Genome Center.  
TITLE Direct Submission  
JOURNAL Unpublished  
REFERENCE  
2 (bases 1 to 129402)  
DOE Joint Genome Institute.  
AUTHORS Direct Submission  
TITLE Submitted (07-OCT-1999) Production Sequencing Facility, DOE Joint  
JOURNAL Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
REFERENCE  
3 (bases 1 to 129402)  
DOE Joint Genome Institute and Stanford Human Genome Center.  
AUTHORS Direct Submission  
TITLE Submitted (23-OCT-2000) DOE Joint Genome Institute, 2800 Mitchell  
JOURNAL Drive, Walnut Creek, CA 94598, USA  
COMMENT On Oct 23, 2000 this sequence version replaced gi:9211205.  
Draft Sequence Produced by DOE Joint Genome Institute  
www.jgi.doe.gov  
Finishing Completed at Stanford Human Genome Center  
www.shgc.stanford.edu  
Quality: Phrap Quality >=40 99.8% of Sequence;  
Estimated Total Number of Errors is 0.7.  
FEATURES  
source  
1..129402  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="19"  
/clone="CTD-2195B23"  
ORIGIN  
Query Match 43.0%; Score 1265.2; DB 9; Length 129402;  
Best Local Similarity 99.4%; Pred. No. 7.4e-222;  
Matches 1270; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
QY 1592 TCTGCGCTCTCTTACGACCATCTCTCAAGCCTTCTCCCTGGAGAGCCCGTGC 1651  
Db 34813 TCTCTCTCTCTTACCAGGAAGCGTGTCTGCTTGGAGAGCCCTTGGCAAAAGCGGAGC 34872  
QY 1652 TCTTCTCTTCTTACCACCATCTCTCAAGCCTTCTCCCTGGAGAGCCCGTGC 1711  
Db 34873 TCTTCTCTTCTTACCACCATCTCTCAAGCCTTCTCCCTGGAGAGCCCGTGC 34932  
QY 1712 ACACCTGAGCCTCAAGCCACCGTCAGTGGCCTTTTCAACATTTCCCGAGCCTTCCAGC 1771  
Db 34933 ACACCTGAGCCTCAAGCCACCGTCAGTGGCCTTTTCAACATTTCCCGAGCCTTCCAGC 34992  
QY 1772 TGCAAGTCCGTCCCTGCTGACCTTCACTCCACACGACGACCATGATGAAGGAGCAACTT 1831  
Db 34993 TGCAAGTCCGTCCCTGCTGACCTTCACTCCACACGACGACCATGATGAAGGAGCAACTT 35052  
QY 1832 GGAAGTGGTGGTGCACAGGACCGTCTTCCAGCCTTCAACAGTGGGATGAGAGGTTA 1891  
Db 35053 GGAAGTGGTGGTGCACAGGACCGTCTTCCAGCCTTCAACAGTGGGATGAGAGGTTA 35112  
QY 1892 ATGCTCCAGAGTGTACACTGCGAGGACGACCATTTACGCGCTGTCAGTGTGTTTTCGGA 1951







Job time : 11160.1 secs

Db	71554	GGAGTGTGGGTGCGCGGAGCGTCCCTCCAGCTCAACAGTGGGCATGGACGGTTA	71495
Qy	1892	ATGCTCCAGAGTGACACTGACGAGCAGCCACATTTACAGCCTGAGTTGTTTCCGGA	1951
Db	71494	ATGCTCCAGAGTGACACTGACGAGCAGCCACATTTACAGCCTGAGTTGTTTCCGGA	71435
Qy	1952	GTCTGTCCACGGCCACACAGCTCACTTGACTCATGCTGCTAAGATGACACACCGCAC	2011
Db	71434	GTCTGTCCACGGCCACACAGCTCACTTGACTCATGCTGCTAAGATGACACACCGCAC	71375
Qy	2012	CCATACACAACCTAAGGGCCACAAAGCAACTGCTGGGTAGCTTTCCACAGACATAAT	2071
Db	71374	CCATACACAACCTAAGGGCCACAAAGCAACTGCTGGGTAGCTTTCCACAGACATAAT	71315
Qy	2072	ATAGTCCATCTGCAATCACAGCAGCATAGCCAGGTAAACCCAACTCCCTGGATCTGC	2131
Db	71314	ATAGTCCATCTGCAATCACAGCAGCATAGCCAGGTAAACCCAACTCCCTGGATCTGC	71255
Qy	2132	AGCCACACAGTGGGAGTCTGGCTGTCACTTCAAGCCACAGAAAGGCCACACATGTT	2191
Db	71254	AGCCACACAGTGGGAGTCTGGCTGTCACTTCAAGCCACAGAAAGGCCACACATGTT	71195
Qy	2192	CACAGCTCACAGGCGCTCTCAATTCATCGAACTTCTCAGTGTCCCTGCTGGTGCCTG	2251
Db	71194	CACAGCTCACAGGCGCTCTCAATTCATCGAACTTCTCAGTGTCCCTGCTGGTGCCTG	71135
Qy	2252	GCACAGGGAACAGCATGCCCTCCGGGTCATGCCACCCAGAGACTGTGCTGTCTATG	2311
Db	71134	GCACAGGGAACAGCATGCCCTCCGGGTCATGCCACCCAGAGACTGTGCTGTCTATG	71075
Qy	2312	GCACCAACTCATGTCCCTCTCTTGGGCTACACCACTCTCCAGGCTGTGACCCAGATGT	2371
Db	71074	GCACCAACTCATGTCCCTCTCTTGGGCTACACCACTCTCCAGGCTGTGACCCAGATGT	71015
Qy	2372	CCACACACCCCAACCACTTGTCCACAGACTACCCACGTACCCAGTACGCTCGTCCCTCCC	2431
Db	71014	CCACACACCCCAACCACTTGTCCACAGACTACCCACGTACCCAGTACGCTCGTCCCTCCC	70955
Qy	2432	CAGAGTATCTCCACCTGACAGCGCGCCGCCACAGAGGACAGTCCCGAGCCCTCT	2491
Db	70954	CAGAGTATCTCCACCTGACAGCGCGCGCCGCCACAGAGGACAGTCCCGAGCCCTCT	70895
Qy	2492	GCAACTGCAGCCCTCAGTCAACCCCTTTTAAAGCACCCCTGATTTACCAATGCAACACA	2551
Db	70894	GCAACTGCAGCCCTCAGTCAACCCCTTTTAAAGCACCCCTGATTTACCAATGCAACACA	70835
Qy	2552	TCTGGGTCTGCGATTATGCAAGAGACTTTGGACATACAGGAGCCCTCAGACCGGAGAA	2611
Db	70834	TCTGGGTCTGCGATTATGCAAGAGACTTTGGACATACAGGAGCCCTCAGACCGGAGAA	70775
Qy	2612	CACCTGCCAACCCCAACAGCTTATGTAACCGTGAAGAGCGGCCCTGCTGCCCC	2671
Db	70774	CACCTGCCAACCCCAACAGCTTATGTAACCGTGAAGAGCGGCCCTGCTGCCCC	70715
Qy	2672	TCCACACACATACACACTCACTGATCTACAGCCCTGTTCCGGCTGAGGTCCCCACT	2731
Db	70714	TCCACACACATACACACTCACTGATCTACAGCCCTGTTCCGGCTGAGGTCCCCACT	70655
Qy	2732	AGACCCAGTGAAGGGTTAGACACCAAGTAGGGCCAGTTTCCAAATCACCCCTGTGAG	2791
Db	70654	AGACCCAGTGAAGGGTTAGACACCAAGTAGGGCCAGTTTCCAAATCACCCCTGTGAG	70595
Qy	2792	GAGTGAGCGGATCTCACGTTCCCTGCTGACTTAAGGGTCCGGCTGGGAATTAAGTTTG	2851
Db	70594	GAGTGAGCGGATCTCACGTTCCCTGCTGACTTAAGGGTCCGGCTGGGAATTAAGTTTG	70535
Qy	2852	TTTCTGGCCTTTAGCCTA	2869
Db	70534	TTTCTGGCCTTTAGCCTA	70517



**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 17, 2004, 06:15:30 ; Search time 1086.48 Seconds  
(without alignments)  
11511.184 Million cell updates/sec

Title: US-10-669-693-1

Perfect score: 2944

Sequence: 1 tttctctgttgcttactc.....aaaaaaaaaaaaaaaaaaaa 2944

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N Geneseq\_29Jan04:\*  
1: Geneseqn1980s:\*  
2: Geneseqn1990s:\*  
3: Geneseqn2000s:\*  
4: Geneseqn2001as:\*  
5: Geneseqn2001bs:\*  
6: Geneseqn2002s:\*  
7: Geneseqn2003as:\*  
8: Geneseqn2003bs:\*  
9: Geneseqn2003cs:\*  
10: Geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	2944	100.0	2944	6	AAD45438
2	2621.4	89.0	2623	4	AAS21275 Human CDN
3	2621.4	89.0	2623	7	ACD23884 Novel hum
4	2621.4	89.0	2623	7	ACA67025 cDNA enco
5	2621.4	89.0	2623	7	ACA03634 cDNA enco
6	2621.4	89.0	2623	7	ABX89172 DNA enco
7	2621.4	89.0	2623	7	ACD41826 Human sec
8	2621.4	89.0	2623	7	ACA04055 Human CDN
9	2621.4	89.0	2623	8	ADA45582 Novel hum
10	2621.4	89.0	2623	8	ADA76013 Human PRO
11	2621.4	89.0	2623	8	ADA18663 Human PRO
12	2621.4	89.0	2623	8	ADA61286 Homo sapi
13	2621.4	89.0	2623	8	ADB19071 Novel hum
14	2621.4	89.0	2623	8	ADB27612 cDNA enco
15	2621.4	89.0	2623	8	ADA86091 Novel hum
16	2621.4	89.0	2623	8	ADB15655 Human PRO
17	2621.4	89.0	2623	8	ADA47441 Human PRO
18	2621.4	89.0	2623	8	ADA67236 Human PRO
19	2621.4	89.0	2623	8	ADB30243 cDNA enco
20	2621.4	89.0	2623	8	ADA85539 Novel hum
21	2621.4	89.0	2623	8	ADA96751 Human PRO
22	2621.4	89.0	2623	8	ADA79055 Human PRO
23	2621.4	89.0	2623	8	ADA87194 Novel hum

24	2621.4	89.0	2623	8	ADB16396	Human PRO
25	2621.4	89.0	2623	8	ADA91488	Novel hum
26	2621.4	89.0	2623	8	ADB14551	Human PRO
27	2621.4	89.0	2623	8	ADB18512	Novel hum
28	2621.4	89.0	2623	8	ADA93727	Human PRO
29	2621.4	89.0	2623	8	ADB19623	Novel hum
30	2621.4	89.0	2623	8	ADB12935	Human PRO
31	2621.4	89.0	2623	8	ACD98455	Novel hum
32	2621.4	89.0	2623	8	ADA74189	Human PRO
33	2621.4	89.0	2623	8	ADB24422	Human PRO
34	2621.4	89.0	2623	8	ADA81946	Human PRO
35	2621.4	89.0	2623	8	ADA74909	Human PRO
36	2621.4	89.0	2623	8	ADA84987	Novel hum
37	2621.4	89.0	2623	8	ADA84435	Novel hum
38	2621.4	89.0	2623	8	ADB29691	cDNA enco
39	2621.4	89.0	2623	8	ADA80219	Human PRO
40	2621.4	89.0	2623	8	ADA75461	Human PRO
41	2621.4	89.0	2623	8	ADA46686	Human PRO
42	2621.4	89.0	2623	8	ADB24982	Human PRO
43	2621.4	89.0	2623	8	ADA93158	Human PRO
44	2621.4	89.0	2623	8	ADB26508	cDNA enco
45	2621.4	89.0	2623	8	ADB30795	cDNA enco

## ALIGNMENTS

RESULT 1  
AAD45438  
ID AAD45438 standard; cDNA; 2944 BP.  
XX  
AC AAD45438;  
XX  
DT 27-DEC-2002 (first entry)  
XX  
DE Human drug-metabolising enzyme encoding cDNA.  
XX  
KW Human; drug-metabolising enzyme; cytochrome p450; therapy; cytostatic;  
KW Parkinson's disease; myotonic dystrophy; developmental defect; enzyme;  
KW cancer; chromosome 19; gene; ss.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT 5'UTR 1..303  
FT /\*tag= a  
FT CDS 304..1818  
FT /\*tag= b  
FT FT /product= "Human drug-metabolising enzyme"  
FT 3'UTR 1819..2944  
FT /\*tag= c  
XX  
US2002076774-A1.  
XX  
PN 20-JUN-2002.  
XX  
PD 27-DEC-2000; 2000US-00748127.  
XX  
PF 21-JUN-2000; 2000US-0212840P.  
XX  
PA (YANG/) YAN C.  
XX (SHAO/) SHAO W.  
XX (MERK/) MERKULOV G V.  
XX (DFRA/) DI FRANCESCO V.  
XX (BEAS/) BEASLEY E M.  
XX  
PI Yan C, Shao W, Merkulov GV, Di Francesco V, Beasley EM;  
XX  
DR WPI; 2002-673347/72.  
XX P-PSDB; AAE28282.  
XX  
PT A new human drug-metabolizing enzyme is related to the cytochrome p450  
PT superfamily and is useful to provide modulators to treat drug-

PT metabolizing enzyme-related disorders including cancer and Parkinson's  
PT disease.  
XX  
PS Claim 4; Page 20-22; 49pp; English.  
XX  
CC The invention relates to human drug-metabolising enzyme peptides related  
CC to the cytochrome p450 superfamily. Binding agents of the peptide are  
CC used to treat a disease or condition mediated by a human drug-  
CC metabolising enzyme. Such diseases include Parkinson's disease, cancer,  
CC myotonic dystrophy and developmental defects. The present sequence is  
CC human drug-metabolising enzyme encoding cDNA. The gene encoding this  
CC enzyme is located at chromosome 19  
XX  
SQ Sequence 2944 BP; 648 A; 957 C; 794 G; 545 T; 0 U; 0 Other;  
Query Match 100.0%; Score 2944; DB 6; Length 2944;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2944; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TTTCTTCTGTTGCTTACTCCCTATCCGGGGGCCCAAGGCGCTGCTCGCGCCCAAGC 60  
DB 1 TTTCTTCTGTTGCTTACTCCCTATCCGGGGGCCCAAGGCGCTGCTCGCGCCCAAGC 60  
QY 61 CCCGGTAAACCTGGGTGACCTCGAGACATCCGTTGGAGCATGAGTCCCGACATCAGG 120  
DB 61 CCCGGTAAACCTGGGTGACCTCGAGACATCCGTTGGAGCATGAGTCCCGACATCAGG 120  
QY 121 CGCGCGCGGTGTCCGGGAGAAACCCGGCGGGGAGAGATAAGCTGCCAGAGGCGAGG 180  
DB 121 CGCGCGCGGTGTCCGGGAGAAACCCGGCGGGGAGAGATAAGCTGCCAGAGGCGAGG 180  
QY 181 GGGCTGGGTAGCTGCCCGCCCGCCGCTGACTTGTGTTGGAGGAGAGACCGCCGGCTC 240  
DB 181 GGGCTGGGTAGCTGCCCGCCCGCCGCTGACTTGTGTTGGAGGAGAGACCGCCGGCTC 240  
QY 241 CGCGCCCTAACTAGCCACCGCGGAGCGCTGGGAGAGAGAGAGAGAGAGAGAGAGAG 300  
DB 241 CGCGCCCTAACTAGCCACCGCGGAGCGCTGGGAGAGAGAGAGAGAGAGAGAGAGAG 300  
QY 301 GAGATGGAGCGGACCGGACCTGGGCGCTGCTGCTGGCGCTGCTGCTGCTGCTGCTG 360  
DB 301 GAGATGGAGCGGACCGGACCTGGGCGCTGCTGCTGGCGCTGCTGCTGCTGCTGCTG 360  
QY 361 AGCTGGCGCTGTCGGGACCGGAGCGGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 420  
DB 361 AGCTGGCGCTGTCGGGACCGGAGCGGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 420  
QY 421 CCACCTGCTGGGAAACCTCCTGAGCTACGGCCCGGGGCGCTGTATTCAGGGCTCATGCGG 480  
DB 421 CCACCTGCTGGGAAACCTCCTGAGCTACGGCCCGGGGCGCTGTATTCAGGGCTCATGCGG 480  
QY 481 CTGAGTAAGAAATAACGAGCGGTGTTCAACATCTACCTGGGACCTGCGGCGCTGCTG 540  
DB 481 CTGAGTAAGAAATAACGAGCGGTGTTCAACATCTACCTGGGACCTGCGGCGCTGCTG 540  
QY 541 GTCTGTTGGGAGAGGCTGCGGAGGCGGCGGAGGCGGAGGCTGAGGAGGTTTCAGG 600  
DB 541 GTCTGTTGGGAGAGGCTGCGGAGGCGGCGGAGGCGGAGGCTGAGGAGGTTTCAGG 600  
QY 601 GGCCTGGGAAACCTGAGCTGCTGGGAGGAGCTTTTGTGGCCATGGGGTTTCTTCTCC 660  
DB 601 GGCCTGGGAAACCTGAGCTGCTGGGAGGAGCTTTTGTGGCCATGGGGTTTCTTCTCC 660  
QY 661 AACGGGAGCGGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 720  
DB 661 AACGGGAGCGGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 720  
QY 721 ATGGGGAAGCGAAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 780  
DB 721 ATGGGGAAGCGAAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 780  
QY 781 TTCCAGGGGACAGAGGACCGCCATTGATTCCTCCTGCTGCTGCTGCTGCTGCTGCTGCT 840

DB 781 TTCCAGGGGACAGAGGAGCGCCATTGATTCCTCCTGCTGCTGCTGCTGCTGCTGCTGCT 840  
QY 841 AACGTAGTCTGCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTC 900  
DB 841 AACGTAGTCTGCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTC 900  
QY 901 GCCGTGGTCCGGGACAGCTGGTGGTACCCTGCTGGGAGTCAGCTCCAGGGGGGTGAGACC 960  
DB 901 GCCGTGGTCCGGGACAGCTGGTGGTACCCTGCTGGGAGTCAGCTCCAGGGGGGTGAGACC 960  
QY 961 TACGAGATGTTCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020  
DB 961 TACGAGATGTTCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020  
QY 1021 CACGTGACACCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080  
DB 1021 CACGTGACACCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080  
QY 1081 GATGCTTCGGGGCCCGCACGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 1140  
DB 1081 GATGCTTCGGGGCCCGCACGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 1140  
QY 1141 GAACAAAACCCAGGACACAGATTCACCAAGAAACATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200  
DB 1141 GAACAAAACCCAGGACACAGATTCACCAAGAAACATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200  
QY 1201 CTGTTTGTGGGACGATGACGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 1260  
DB 1201 CTGTTTGTGGGACGATGACGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 1260  
QY 1261 AAATACCTCTCATGTCCAAAAGTGGGTACGTGAGGAGCTGAAATCGGGAGCTGGGGGTGCG 1320  
DB 1261 AAATACCTCTCATGTCCAAAAGTGGGTACGTGAGGAGCTGAAATCGGGAGCTGGGGGTGCG 1320  
QY 1321 CAGGACCAAGCTGAGGGACCGTACCGCTCCCTTACACCGACGCGGTTCGTCATGAG 1380  
DB 1321 CAGGACCAAGCTGAGGGACCGTACCGCTCCCTTACACCGACGCGGTTCGTCATGAG 1380  
QY 1381 GCGGACCGGCTGCTGGCGTGGTCCCATGGGAATACCCGCGACCTCATGGGAGCACACC 1440  
DB 1381 GCGGACCGGCTGCTGGCGTGGTCCCATGGGAATACCCGCGACCTCATGGGAGCACACC 1440  
QY 1441 CGCTTCCAGAGGTACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500  
DB 1441 CGCTTCCAGAGGTACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500  
QY 1501 CTGCTATGACCCCAACATCTTCAAGCAACCAAGAGAGTTCACCCAGACCGTTTCTGGAT 1560  
DB 1501 CTGCTATGACCCCAACATCTTCAAGCAACCAAGAGAGTTCACCCAGACCGTTTCTGGAT 1560  
QY 1561 GCAGATGAGCGGTTGAGGAAGCATGAGGGGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1620  
DB 1561 GCAGATGAGCGGTTGAGGAAGCATGAGGGGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1620  
QY 1621 TGCTTGGAGAGGGCTGGCAAAAGCGAGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1680  
DB 1621 TGCTTGGAGAGGGCTGGCAAAAGCGAGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1680  
QY 1681 GCCTTCTCCTCGGAGAGCCGTCGCGCGCGGACACCTGAGGCTCAAGCCACCGTCAGT 1740  
DB 1681 GCCTTCTCCTCGGAGAGCCGTCGCGCGCGGACACCTGAGGCTCAAGCCACCGTCAGT 1740  
QY 1741 GGCCTTTTCAACATTTCCCGGCTTCCAGCTGCAAGTCCGCTCCCACTGACCTTCACTCC 1800  
DB 1741 GGCCTTTTCAACATTTCCCGGCTTCCAGCTGCAAGTCCGCTCCCACTGACCTTCACTCC 1800  
QY 1801 ACCAGCAGACAGATGAAAGGAGGCACTTGGAGTGTGGGTGCGGAGGAGCGTGCCT 1860  
DB 1801 ACCAGCAGACAGATGAAAGGAGGCACTTGGAGTGTGGGTGCGGAGGAGCGTGCCT 1860  
QY 1861 CCAGCTCAACAGTGGGATGAGAGGGTTAATGCTCCAGAGTGTACACTGCGAGGCGAGC 1920  
DB 1861 CCAGCTCAACAGTGGGATGAGAGGGTTAATGCTCCAGAGTGTACACTGCGAGGCGAGC 1920

1921 QY CACATTTACACGCTGCAGTTGTTTTCCGGAGTCTGTCACACGCGCCACACGCTCACTTG 1980  
1921 Db CACATTTACACGCTGCAGTTGTTTTCCGGAGTCTGTCACACGCGCCACACGCTCACTTG 1980  
1981 QY ACTCATGCTGCTAAGATGCAACACCGCACACCCATACACAACTACAAGGGCCACAAAGCA 2040  
1981 Db ACTCATGCTGCTAAGATGCAACACCGCACACCCATACACAACTACAAGGGCCACAAAGCA 2040  
2041 QY ACTGCTGGGTAGTTTCCACAGACATAAATATAGTCCATCTGCAATCAAAAGCACATAG 2100  
2041 Db ACTGCTGGGTAGTTTCCACAGACATAAATATAGTCCATCTGCAATCAAAAGCACATAG 2100  
2101 QY CCAAGTAACCCCAACTCCCTCGATCTGCAGCCACACAGTGGGATCTGGGTGTCAAC 2160  
2101 Db CCAAGTAACCCCAACTCCCTCGATCTGCAGCCACACAGTGGGATCTGGGTGTCAAC 2160  
2161 QY TTCACAAGCCACAGAAACGCGCCACACATGTTTCACAGCTCACAGCCCTCTCCCAATTCATCG 2220  
2161 Db TTCACAAGCCACAGAAACGCGCCACACATGTTTCACAGCTCACAGCCCTCTCCCAATTCATCG 2220  
2221 QY AACTTCTCAGTGTCCCTGTCCTGTCCTGGTGGCCACAGGGAACAGCATGCCCCCTCCGGGG 2280  
2221 Db AACTTCTCAGTGTCCCTGTCCTGTCCTGGTGGCCACAGGGAACAGCATGCCCCCTCCGGGG 2280  
2281 QY TCATGCGCCACGAGACTGTCGCTGCTATGCGCCCAACTCATGCTCCCTCTCTGGCTA 2340  
2281 Db TCATGCGCCACGAGACTGTCGCTGCTATGCGCCCAACTCATGCTCCCTCTCTGGCTA 2340  
2341 QY CACCACTCTCCAGCCTGTGACCAACCCGATGTCACACACACCCCAACACATGTTGCCACACA 2400  
2341 Db CACCACTCTCCAGCCTGTGACCAACCCGATGTCACACACACCCCAACACATGTTGCCACACA 2400  
2401 QY GCTTACCCAGTACGACATGCTCTGCTGCTCCCGAGAGTATCTTCCCATGAGACACGCCGC 2460  
2401 Db GCTTACCCAGTACGACATGCTCTGCTGCTCCCGAGAGTATCTTCCCATGAGACACGCCGC 2460  
2461 QY CCCCACAGAGGACAGTCCCGAGCCACTCTGCAACTGCGAGCCCTCAGTCAACCCCTTTT 2520  
2461 Db CCCCACAGAGGACAGTCCCGAGCCACTCTGCAACTGCGAGCCCTCAGTCAACCCCTTTT 2520  
2521 QY AAGCACCTCTGATTCTACCAATGCAACACATCTGGGCTCTGCGATTTATGACAGAGACTT 2580  
2521 Db AAGCACCTCTGATTCTACCAATGCAACACATCTGGGCTCTGCGATTTATGACAGAGACTT 2580  
2581 QY TGACATACAGAGACCTCTGACCGGAGGAACACCTGCGCCCAACCCCAACACGTTGCTTATG 2640  
2581 Db TGACATACAGAGACCTCTGACCGGAGGAACACCTGCGCCCAACCCCAACACGTTGCTTATG 2640  
2641 QY TAACACGTTGGAAGCGGCCCTGCTGCCCTCCACACACACATACACACTCACTGATCT 2700  
2641 Db TAACACGTTGGAAGCGGCCCTGCTGCCCTCCACACACACATACACACTCACTGATCT 2700  
2701 QY ACAGCCCTCTGTCGGGCTCAGAGTCCCACTAGACCCAGTGGAGGGGTTAGAGACCAAG 2760  
2701 Db ACAGCCCTCTGTCGGGCTCAGAGTCCCACTAGACCCAGTGGAGGGGTTAGAGACCAAG 2760  
2761 QY TAGGGGCGAGTTTCCAAATTCACCTGTGAGGAGTGAGCGGATCTGACGTTTCTTGTGA 2820  
2761 Db TAGGGGCGAGTTTCCAAATTCACCTGTGAGGAGTGAGCGGATCTGACGTTTCTTGTGA 2820  
2821 QY CTTTAAAGGTCGGGCTTGGGAATTAAGTTTGTCTGCGCTTTAGCTTAAAAA 2880  
2821 Db CTTTAAAGGTCGGGCTTGGGAATTAAGTTTGTCTGCGCTTTAGCTTAAAAA 2880  
2881 QY AA 2940  
2881 Db AA 2940  
2941 QY AAAAA 2944  
2941 Db AAAAA 2944

RESULT 2  
AAS21275  
ID AAS21275 standard; cDNA; 2623 BP.  
XX  
AC AAS21275;  
DT 24-OCT-2001 (first entry)  
XX  
Human cDNA sequence encoding for PRO1906 polypeptide.  
DE Human secretory and transmembrane; PRO; mammalian; cancer; lung; breast;  
XX prostate; cervical; tumour necrosis factor-alpha; TNF-alpha; cartilage;  
KW ear; proliferation; glucose; free fatty acid; skeletal muscle; adipocyte;  
KW A-peptide; factor VIIA; gene therapy; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200140466-A2.  
XX  
PD 07-JUN-2001.  
XX  
PF 01-DEC-2000; 2000WO-US032678.  
XX  
PR 01-DEC-1999; 99WO-US028301.  
PR 01-DEC-1999; 99WO-US028634.  
PR 02-DEC-1999; 99WO-US028551.  
PR 02-DEC-1999; 99WO-US028564.  
PR 02-DEC-1999; 99WO-US028565.  
PR 09-DEC-1999; 99US-0170262P.  
PR 16-DEC-1999; 99WO-US030095.  
PR 20-DEC-1999; 99WO-US030911.  
PR 20-DEC-1999; 99WO-US030999.  
PR 30-DEC-1999; 99WO-US031243.  
PR 30-DEC-1999; 99WO-US031274.  
PR 05-JAN-2000; 2000WO-US000219.  
PR 06-JAN-2000; 2000WO-US000277.  
PR 06-JAN-2000; 2000WO-US000376.  
PR 11-FEB-2000; 2000WO-US003565.  
PR 18-FEB-2000; 2000WO-US004341.  
PR 18-FEB-2000; 2000WO-US004342.  
PR 22-FEB-2000; 2000WO-US004414.  
PR 24-FEB-2000; 2000WO-US004914.  
PR 24-FEB-2000; 2000WO-US005004.  
PR 01-MAR-2000; 2000WO-US005601.  
PR 02-MAR-2000; 2000WO-US005841.  
PR 03-MAR-2000; 2000US-0187202P.  
PR 10-MAR-2000; 2000WO-US006319.  
PR 15-MAR-2000; 2000WO-US006884.  
PR 20-MAR-2000; 2000WO-US007377.  
PR 21-MAR-2000; 2000WO-US007532.  
PR 30-MAR-2000; 2000WO-US008439.  
PR 17-MAY-2000; 2000WO-US013705.  
PR 22-MAY-2000; 2000WO-US014042.  
PR 30-MAY-2000; 2000WO-US014941.  
PR 02-JUN-2000; 2000WO-US015264.  
PR 05-JUN-2000; 2000US-0209832P.  
PR 28-JUL-2000; 2000WO-US020710.  
PR 11-AUG-2000; 2000WO-US020231.  
PR 23-AUG-2000; 2000WO-US023522.  
PR 24-AUG-2000; 2000WO-US023328.  
PR 08-NOV-2000; 2000WO-US030952.  
PR 10-NOV-2000; 2000WO-US030873.  
XX  
(GETH ) GENENTECH INC.  
XX  
XX Baker KP, Beresini M, Deforge L, Desnoyers L, Filvaroff E, Gao W;  
PI Gerritsen ME, Goddard A, Godowski PU, Gurney AL, Sherwood S;  
PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;  
XX  
XX WPI; 2001-408281/43.  
DR P-PSDB; AAU12203.  
XX



QY 1944 TTTCGGAGTCTGTCACAGCCGCCACAGCTCACTTGACTCATGCTGCTAGATGACAA 2003  
DB 1581 TTTCGGAGTCTGTCACAGCCGCCACAGCTCACTTGACTCATGCTGCTAGATGACAA 1740  
QY 2004 CCGCACACCCATACACAACCTACAAGGGCCACAAGGCACTCTGGTTAGCTTTCCACAG 2063  
DB 1741 CCGCACACCCATACACAACCTACAAGGGCCACAAGGCACTCTGGTTAGCTTTCCACAG 1800  
QY 2064 ACATAAATAGTCCATCTGCAATCAAGACACATAGCCAGGTAAACCAACATCCGCT 2123  
DB 1801 ACATAAATAGTCCATCTGCAATCAAGACACATAGCCAGGTAAACCAACATCCGCT 1860  
QY 2124 GGATCTGCAGCCACACAGTGGAGTCTGGCTGTCACTTCAAGCCACACAGAAAGGCCA 2183  
DB 1861 GGATCTGCAGCCACACAGTGGAGTCTGGCTGTCACTTCAAGCCACACAGAAAGGCCA 1920  
QY 2184 CACATGTTCAAGTCTCAAGCCCTCTCCATTCATCGAACTTCTCAGTGTCCCTGTCCCT 2243  
DB 1921 CACATGTTCAAGTCTCAAGCCCTCTCCATTCATCGAACTTCTCAGTGTCCCTGTCCCT 1980  
QY 2244 GGTGCTGGCACAGGGAACAGTATGCCCTCCGGGTCTATGCCACCCAGAGACTGTGCG 2303  
DB 1981 GGTGCTGGCACAGGGAACAGTATGCCCTCCGGGTCTATGCCACCCAGAGACTGTGCG 2040  
QY 2304 TGTCTATGGCCCAACTCATGCTCCCTCTCTTGCTACACCACTCTCCAGCCCTGTGACC 2363  
DB 2041 TGTCTATGGCCCAACTCATGCTCCCTCTCTTGCTACACCACTCTCCAGCCCTGTGACC 2100  
QY 2364 ACGATGTTCCACACACCCCAACCACTTGTTCCACACAGCTACCCACGTAACGACATCGTCC 2423  
DB 2101 ACGATGTTCCACACACCCCAACCACTTGTTCCACACAGCTACCCACGTAACGACATCGTCC 2160  
QY 2424 TGGTCTCCCAAGAGTATCTTCCACTGACACAGCGCGGCCCCACAGAGGACAGTCCGAG 2483  
DB 2161 TGGTCTCCCAAGAGTATCTTCCACTGACACAGCGCGGCCCCACAGAGGACAGTCCGAG 2220  
QY 2484 CCACCTCTGCAACTGACGCTCAGTCAACCCCTTTTAAAGCACCTGATTTACCAATG 2543  
DB 2221 CCACCTCTGCAACTGACGCTCAGTCAACCCCTTTTAAAGCACCTGATTTACCAATG 2280  
QY 2544 CAAACACATCTGGTCTGCGATTATGACAGAGACTTTGGACATACAGGACCCCTCAGAC 2603  
DB 2281 CAAACACATCTGGTCTGCGATTATGACAGAGACTTTGGACATACAGGACCCCTCAGAC 2340  
QY 2604 CGGAGGAACCTGCCCAACCCCAACAGCTGCTTATGTAACCAAGTGGAAAGCGGCCCT 2663  
DB 2341 CGGAGGAACCTGCCCAACCCCAACAGCTGCTTATGTAACCAAGTGGAAAGCGGCCCT 2400  
QY 2664 GCTGCCCTCCACACACATACACACTCACTGATCTACAGCCCTGTTGCGGCTCAGAG 2723  
DB 2401 GCTGCCCTCCACACACATACACACTCACTGATCTACAGCCCTGTTGCGGCTCAGAG 2460  
QY 2724 TCCCCACTAGACCCAGTGGAGGGTTAGAGACCAAGTAGGGCCAGTTTCCAAATTCACC 2783  
DB 2461 TCCCCACTAGACCCAGTGGAGGGTTAGAGACCAAGTAGGGCCAGTTTCCAAATTCACC 2520  
QY 2784 CTGTCAGGAGTGAAGCGGATCTGACGCTCTGTTGACTTAAGGGTCCGCTTGGGAATT 2843  
DB 2521 CTGTCAGGAGTGAAGCGGATCTGACGCTCTGTTGACTTAAGGGTCCGCTTGGGAATT 2580  
QY 2844 AAAGTTTGTCTTGGCCCTTTAGCCTAAAAAAGGAAAAA 2886  
DB 2581 AAAGTTTGTCTTGGCCCTTTAGCCTAAAAAAGGAAAAA 2623

## RESULT 3

ACD23884

ID ACD23884 standard; cDNA; 2623 BP.

XX

AC ACD23884;

XX

DT 26-AUG-2003 (first entry)

XX

DE Novel human secreted and transmembrane protein PRO1906 cDNA.  
XX Human; secreted and transmembrane protein; PRO; antiinflammatory;  
KW antiarteriosclerotic; cardiant; anti-infertility; anti-HIV; cytostatic;  
KW antidiabetic; gene therapy; tumour necrosis factor (TNF)-alpha release;  
KW TNF-alpha release; cell proliferation; cell differentiation;  
KW gene expression modulator; proteoglycan release; cytokine release;  
KW tumour; inflammatory disease; organ failure; atherosclerosis;  
KW cardiac injury; infertility; birth defect; premature aging; AIDS;  
KW acquired immunodeficiency syndrome; cancer; diabetic complication;  
KW chromosome mapping; gene mapping; pharmaceutical; diagnostic; biosensor;  
KW bioreactor; tissue typing; gene; ss.

Homo sapiens.

US2003032156-A1.

13-FEB-2003.

06-MAY-2002; 2002US-00140474.

31-MAR-1997; 97WO-US005230.

12-JUN-1998; 98WO-US012456.

14-JUL-1998; 98WO-US014552.

28-AUG-1998; 98WO-US017888.

10-SEP-1998; 98WO-US018824.

14-SEP-1998; 98WO-US019093.

14-SEP-1998; 98WO-US019094.

16-SEP-1998; 98WO-US019177.

17-SEP-1998; 98WO-US019330.

07-OCT-1998; 98WO-US019437.

29-OCT-1998; 98WO-US021141.

29-OCT-1998; 98WO-US022991.

29-OCT-1998; 98WO-US022992.

20-NOV-1998; 98WO-US024855.

01-DEC-1998; 98WO-US025108.

05-JAN-1999; 98WO-US000106.

08-MAR-1999; 99WO-US0005028.

10-MAR-1999; 99WO-US0005190.

20-APR-1999; 99WO-US008615.

14-MAY-1999; 99WO-US010733.

02-JUN-1999; 99WO-US012252.

01-SEP-1999; 99WO-US020111.

08-SEP-1999; 99WO-US020594.

13-SEP-1999; 99WO-US020944.

15-SEP-1999; 99WO-US021090.

15-SEP-1999; 99WO-US021547.

05-OCT-1999; 99WO-US023089.

29-NOV-1999; 99WO-US028214.

30-NOV-1999; 99WO-US028313.

01-DEC-1999; 99WO-US028409.

01-DEC-1999; 99WO-US028301.

01-DEC-1999; 99WO-US028634.

02-DEC-1999; 99WO-US028551.

02-DEC-1999; 99WO-US028564.

02-DEC-1999; 99WO-US028565.

16-DEC-1999; 99WO-US030095.

20-DEC-1999; 99WO-US030911.

20-DEC-1999; 99WO-US030999.

22-DEC-1999; 99WO-US030720.

30-DEC-1999; 99WO-US031243.

30-DEC-1999; 99WO-US031274.

05-JAN-2000; 2000WO-US000219.

06-JAN-2000; 2000WO-US000277.

06-JAN-2000; 2000WO-US000376.

11-FEB-2000; 2000WO-US003565.

18-FEB-2000; 2000WO-US004341.

18-FEB-2000; 2000WO-US004342.

22-FEB-2000; 2000WO-US004414.

24-FEB-2000; 2000WO-US004914.

24-FEB-2000; 2000WO-US005004.

01-MAR-2000; 2000WO-US005601.

02-MAR-2000; 2000WO-US005746.

PR 02-MAR-2000; 2000WO-US005841.  
PR 10-MAR-2000; 2000WO-US006319.  
PR 15-MAR-2000; 2000WO-US006884.  
PR 20-MAR-2000; 2000WO-US007377.  
PR 21-MAR-2000; 2000WO-US007532.  
PR 30-MAR-2000; 2000WO-US008439.  
PR 17-MAY-2000; 2000WO-US013705.  
PR 22-MAY-2000; 2000WO-US014042.  
PR 30-MAY-2000; 2000WO-US014941.  
PR 02-JUN-2000; 2000WO-US015264.  
PR 28-JUL-2000; 2000WO-US020710.  
PR 11-AUG-2000; 2000WO-US022031.  
PR 23-AUG-2000; 2000WO-US023522.  
PR 24-AUG-2000; 2000WO-US023328.  
PR 08-NOV-2000; 2000WO-US030952.  
PR 10-NOV-2000; 2000WO-US030873.  
PR 01-DEC-2000; 2000WO-US032678.  
PR 20-DEC-2000; 2000US-00747259.  
PR 20-DEC-2000; 2000WO-US034956.  
PR 28-FEB-2001; 2001US-00796498.  
PR 28-FEB-2001; 2001WO-US006520.  
PR 01-MAR-2001; 2001WO-US006666.  
PR 09-MAR-2001; 2001US-00802706.  
PR 14-MAR-2001; 2001US-00808689.  
PR 22-MAR-2001; 2001US-00816744.  
PR 05-APR-2001; 2001US-00828366.  
PR 10-MAY-2001; 2001US-00854208.  
PR 10-MAY-2001; 2001US-00854280.  
PR 18-MAY-2001; 2001US-00860216.  
PR 25-MAY-2001; 2001US-00866028.  
PR 25-MAY-2001; 2001US-00866034.  
PR 25-MAY-2001; 2001WO-US017092.  
PR 01-JUN-2001; 2001US-00872035.  
PR 01-JUN-2001; 2001WO-US017800.  
PR 05-JUN-2001; 2001US-00874503.  
PR 14-JUN-2001; 2001US-00882636.  
PR 19-JUN-2001; 2001US-00886342.  
PR 20-JUN-2001; 2001WO-US019692.  
PR 21-JUN-2001; 2001US-00887879.  
PR 22-JUN-2001; 2001WO-US020116.  
PR 29-JUN-2001; 2001WO-US021066.  
PR 09-JUL-2001; 2001WO-US021735.  
PR 18-JUL-2001; 2001US-00908827.  
PR 06-AUG-2001; 2001US-00924419.  
PR 09-AUG-2001; 2001US-00927796.  
PR 16-AUG-2001; 2001US-00931836.  
PR 19-DEC-2001; 2001US-00028072.

XX (GETH ) GENENTECH INC.

XX Baker KP, Beresini M, Deforge L, Deenoyers L, Filvaroff E, Gao W;  
PI Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;  
PI Smith V, Stewart IA, Tumas D, Watanabe CK, Wood WI, Zhang Z;  
XX WPI; 2003-341980/32.  
DR P-PSDB; AB017647.

XX New secreted and transmembrane PRO nucleic acids, for treating  
PT inflammation, organ failure, atherosclerosis, cardiac injury,  
PT infertility, birth defects, premature aging, acquired immunodeficiency  
PT syndrome (AIDS), or cancer.

XX Claim 2; Fig 63; 660pp; English.

XX The invention describes an isolated nucleic acid (I) comprising, or which  
CC has 80 % sequence identity to, or the full-length coding sequence of, one  
CC of 275 nucleotide sequences, and which encodes a corresponding  
CC polypeptide selected from 275 amino acid sequences, where all sequences  
CC are given in the specification. The polypeptide encoded by (I) is used to  
CC detect PRO polypeptides, link a bioactive molecule to a cell expressing a  
CC PRO polypeptide, modulate a biological activity of a cell, stimulate the  
CC release of tumour necrosis factor (TNF)-alpha from human blood, modulate  
CC the uptake of glucose or free fatty acid by cells, stimulate or inhibit

CC the proliferation or differentiation of cells or gene expression,  
CC stimulate the release of proteoglycans, stimulate the release of cytokine  
CC from peripheral blood mononuclear cells, inhibit the binding of A-peptide  
CC to factor VIIA, or detect the presence of tumour in a mammal. The nucleic  
CC acid and polypeptide encoded by it, are useful for treating inflammatory  
CC diseases, organ failure, atherosclerosis, cardiac injury, infertility,  
CC birth defects, premature aging, acquired immunodeficiency syndrome  
CC (AIDS), cancer, or diabetic complications. The nucleic acid is useful as  
CC hybridisation probes, in chromosome and gene mapping, and in generating  
CC antisense RNA or DNA. The polypeptides are useful as pharmaceuticals,  
CC diagnostics, biosensors or bioreactors. Both are useful in tissue typing.  
CC This sequence encodes a novel human secreted and transmembrane PRO  
CC polypeptide

XX Sequence 2623 BP; 552 A; 865 C; 704 G; 502 T; 0 U; 0 Other;

Query Match 89.0%; Score 2621.4; DB 7; Length 2623;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2622; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 264 GCGGAGCGCTCGGAGAGGAGGAGCGGAGCTGCCGAGATGGAGCGGACCGGACCTTG 323  
DB 1 GCGGAGCGCTCGGAGAGGAGGAGCGGAGCTGCCGAGATGGAGCGGACCGGACCTTG 60  
QY 324 GCGGCTGTGCTGGCGCTGGCGCTGCTCTCTGTCTGTACGCTGGCGCTGTCCGGGACCA 383  
DB 61 GCGGCTGTGCTGGCGCTGGCGCTGCTCTCTGTCTGTACGCTGGCGCTGTCCGGGACCA 120  
QY 384 GCGGAGGCGACCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 443  
DB 121 GCGGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 180  
QY 444 GCTACGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 503  
DB 181 GCTACGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 240  
QY 504 GTTACCATCTACCTGGGACCTTGGGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 563  
DB 241 GTTACCATCTACCTGGGACCTTGGGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 300  
QY 564 GCGGAGGCGCTGGGAGGTCAGGCTGAGGAGTTTACGCGCGGCGGCGGCGGCGGCGGCGG 623  
DB 301 GCGGAGGCGCTGGGAGGTCAGGCTGAGGAGTTTACGCGCGGCGGCGGCGGCGGCGGCGG 360  
QY 624 GGAAGGGACTTTTGATGGGCTATGGGCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 683  
DB 361 GGAAGGGACTTTTGATGGGCTATGGGCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 420  
QY 684 GAGGAAGTTTACATGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 743  
DB 421 GAGGAAGTTTACATGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 480  
QY 744 GCTGATCCAGGCGGAGGCGGCGGCTGTCTGTGTGAGACATTCACAGGCGGAGGAGGCGG 803  
DB 481 GCTGATCCAGGCGGAGGCGGCGGCTGTCTGTGTGAGACATTCACAGGCGGAGGAGGCGG 540  
QY 804 ATTGATCCCTCCCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 863  
DB 541 ATTGATCCCTCCCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 600  
QY 864 TGGCCTCCGCTTCTCTATGAGATTAAGAGTTTCCAGGCGGCTGTCTGTCTGTCTGTCTGT 923  
DB 601 TGGCCTCCGCTTCTCTATGAGATTAAGAGTTTCCAGGCGGCTGTCTGTCTGTCTGTCTGT 660  
QY 924 TACCCTGTCTGGGAGTCAGCTCCAGGGGGTTCAGACTACGAGATCTTCTCTGTCTCTCT 983  
DB 661 TACCCTGTCTGGGAGTCAGCTCCAGGGGGTTCAGACTACGAGATCTTCTCTGTCTCTCT 720  
QY 984 GCGGCGGCTTCCAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1043  
DB 721 GCGGCGGCTTCCAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 780  
QY 1044 CACAGTCCGGGAGGTGCAGCAGCACCCAGGGGAACTTGGATGCTTCCGGGCGGCGGCGG 1103



Db 781 ||||| CACAGTCCGGCAGGTGACAGCAGCACCGGGAACCTGATGCTTCGGGCCCGCAGCTGA 840  
QY 1104 CTTGTGATGCTCTCTGCTGAGATGGCAAGAGGAAACAAACCCAGGACACAGAATT 1163  
Db 841 CTTGTGATGCTCTCTGCTGAGATGGCAAGAGGAAACAAACCCAGGACACAGAATT 900  
QY 1164 CACAAACAAAGAACATGCTGATGACAGTCAATTATTGCTGTTGCTGGGACGATGACGGT 1223  
Db 901 CACCAACAAAGAACATGCTGATGACAGTCAATTATTGCTGTTGCTGGGACGATGACGGT 960  
QY 1224 CAGCACCAACCGTCCGCTATACCCCTCTGCTGATGAAATACCTCATGTCCAAAAGTG 1283  
Db 961 CAGCACCAACCGTCCGCTATACCCCTCTGCTGATGAAATACCTCATGTCCAAAAGTG 1020  
QY 1284 GGTACGTGAGAGCTGAATCCGGAGCTGGGGCTGGCCAGGACCAAGCCTTAGGGACCG 1343  
Db 1021 GGTACGTGAGAGCTGAATCCGGAGCTGGGGCTGGCCAGGACCAAGCCTTAGGGACCG 1080  
QY 1344 TACCCGCTCCCTTACACCGACGCGTTCTGCATGAGCGCAGCGGCTGCTGGCGTGGT 1403  
Db 1081 TACCCGCTCCCTTACACCGACGCGTTCTGCATGAGCGCAGCGGCTGCTGGCGTGGT 1140  
QY 1404 GCCCATGGGAATACCCCGACCCCTCATGCGGACCAACCGCTTCCGAGGGGTACACCCCTGCC 1463  
Db 1141 GCCCATGGGAATACCCCGACCCCTCATGCGGACCAACCGCTTCCGAGGGGTACACCCCTGCC 1200  
QY 1464 CCAGGACCGAGGCTCTTCCCGCTCTCTGCTGATGAGCGCAGCGGCTGCTGGCGTGGT 1523  
Db 1201 CCAGGACCGAGGCTCTTCCCGCTCTCTGCTGATGAGCGCAGCGGCTGCTGGCGTGGT 1260  
QY 1524 GCACCCAGAGAGTTCAACCCAGACCGTTTCTGATGAGCGCAGCGGCTGCTGGCGTGGT 1583  
Db 1261 GCACCCAGAGAGTTCAACCCAGACCGTTTCTGATGAGCGCAGCGGCTGCTGGCGTGGT 1320  
QY 1584 TGAGGCGTCTCTGCTCTCTCTTAGGGAAGCGTGTCTGCTTGGAGGGCGCTGGCAAA 1643  
Db 1321 TGAGGCGTCTCTGCTCTCTCTTAGGGAAGCGTGTCTGCTTGGAGGGCGCTGGCAAA 1380  
QY 1644 AGCGAGCTCTCTCTCTCTCTTACACCAATCTCTCAAGCGTCTTCCCTGAGAGCGCGTG 1703  
Db 1381 AGCGAGCTCTCTCTCTCTCTTACACCAATCTCTCAAGCGTCTTCCCTGAGAGCGCGTG 1440  
QY 1704 CCGCGCGACACCTGAGCTCAAGCCACCGTCAGTGGCGCTTTCACATTCGCCAGC 1763  
Db 1441 CCGCGCGACACCTGAGCTCAAGCCACCGTCAGTGGCGCTTTCACATTCGCCAGC 1500  
QY 1764 CTTCCAGCTGCAAGTCCGTCCCACTGACCTTCACTCCACCAACGACAGATGAAGGAA 1823  
Db 1501 CTTCCAGCTGCAAGTCCGTCCCACTGACCTTCACTCCACCAACGACAGATGAAGGAA 1560  
QY 1824 GGCAACTTGAAGTGGTGGTGGCCAGGACGGTCTCCAGCGTCAACATGGGGATGGA 1883  
Db 1561 GGCAACTTGAAGTGGTGGTGGCCAGGACGGTCTCCAGCGTCAACATGGGGATGGA 1620  
QY 1884 CAGGGTTAATGCTCTCAGAGTGATACCTGAGCGAGCGCAATTACAGCGCTGAGTTGT 1943  
Db 1621 CAGGGTTAATGCTCTCAGAGTGATACCTGAGCGAGCGCAATTACAGCGCTGAGTTGT 1680  
QY 1944 TTTCCGAGTCTCTCCACCGGCCACACGCTCACTTGACTCATGCTGCTAAGATGACAA 2003  
Db 1681 TTTCCGAGTCTCTCTCCACCGGCCACACGCTCACTTGACTCATGCTGCTAAGATGACAA 1740  
QY 2004 CCGCACCCCATACACAACTACAGGGCCCAAGGAACTGTGGTTAGCTTCCACAG 2063  
Db 1741 CCGCACCCCATACACAACTACAGGGCCCAAGGAACTGTGGTTAGCTTCCACAG 1800  
QY 2064 ACATAAATAGTCCATCTGCAATCAACAGCATAGCCAGGTAAACCCACCACTCCCT 2123  
Db 1801 ACATAAATAGTCCATCTGCAATCAACAGCATAGCCAGGTAAACCCACCACTCCCT 1860  
QY 2124 GGATCTGCAGCCACACGCTGGGAGTCTGGCTGTACCTTCCACAGGACACAGAAACGGCCA 2183

Db 1861 GGATCTGCAGCCCAACAAGTGGAGTCTGGCTGTCACTTCAAGCCACAGAAACGGCCA 1920  
QY 2184 CACATGTTACAGCTCAACAGCCCTCTCTCATTTCAATCGAACTTCTCAGTGTCCCTTCCT 2243  
Db 1921 CACATGTTACAGCTCAACAGCCCTCTCTCATTTCAATCGAACTTCTCAGTGTCCCTTCCT 1980  
QY 2244 GGTGCTGGGACAGGGAACAGATGCCCCCTCCGGGTCATGCCACCCAGAGATGTGCG 2303  
Db 1981 GGTGCTGGGACAGGGAACAGATGCCCCCTCCGGGTCATGCCACCCAGAGATGTGCG 2040  
QY 2304 TGTCTATGGCCCCCAACTCATGCTCCCTCTCTTGGCTACACCACTCTCCAGCCTGTGACC 2363  
Db 2041 TGTCTATGGCCCCCAACTCATGCTCCCTCTCTTGGCTACACCACTCTCCAGCCTGTGACC 2100  
QY 2364 ACCGATGTCCACACACACCCCAACCACTTGTCCACACAGCTACCCACGTAACGACATCGTCC 2423  
Db 2101 ACCGATGTCCACACACACCCCAACCACTTGTCCACACAGCTACCCACGTAACGATCGTCC 2160  
QY 2424 TGGCTCCCGAGAGTATCTTCCACTGAGACAGCGCGCCCCACAGAGGCAAGTCCCCAG 2483  
Db 2161 TGGCTCCCGAGAGTATCTTCCACTGAGACAGCGCGCCCCACAGAGGCAAGTCCCCAG 2220  
QY 2484 CCACCTCTGCAACTGACAGCCTCAGTCAACCCCTTTTAAAGCACCTGATTTACCAATG 2543  
Db 2221 CCACCTCTGCAACTGACAGCCTCAGTCAACCCCTTTTAAAGCACCTGATTTACCAATG 2280  
QY 2544 CAAACATCTGGTCTGCGATTATGACAGAGACTTTTGGACATACGAGGACCTTCAGAC 2603  
Db 2281 CAAACATCTGGTCTGCGATTATGACAGAGACTTTTGGACATACGAGGACCTTCAGAC 2340  
QY 2604 CGGAGGAACCTTCCCAACCCCAACCGTGTATGTAAACCAAGTGGAAAGCGGCCCT 2663  
Db 2341 CGGAGGAACCTTCCCAACCCCAACCGTGTATGTAAACCAAGTGGAAAGCGGCCCT 2400  
QY 2664 GCTGCCCTCCACACACACATACACTCACTCATCTACAGCCCTGTTCCCAATTCACC 2723  
Db 2401 GCTGCCCTCCACACACATACACTCACTCATCTACAGCCCTGTTCCCAATTCACC 2460  
QY 2724 TCCCACTAGACCCAGTGAAGGGTTAGAGACCAAGTAGGGGCCAGTTTCCCAATTCACC 2783  
Db 2461 TCCCACTAGACCCAGTGAAGGGTTAGAGACCAAGTAGGGGCCAGTTTCCCAATTCACC 2520  
QY 2784 CTCTCAGGAGTGAAGCGGATCTGAGCTTCTGTGACCTTAAGGGTCCGCTTGGGAATT 2843  
Db 2521 CTCTCAGGAGTGAAGCGGATCTGAGCTTCTGTGACCTTAAGGGTCCGCTTGGGAATT 2580  
QY 2844 AAAGTTTGTCTTGGCTTTAGCCTTAAAAAATAAAAAA 2886  
Db 2581 AAAGTTTGTCTTGGCTTTAGCCTTAAAAAATAAAAAA 2623

## RESULT 4

ACA67025

ID ACA67025 standard; cDNA; 2623 BP.

XX ACA67025;

AC ACA67025;

XX ACA67025;

DT 23-JUN-2003 (first entry)

XX cDNA encoding human PRO polypeptide #32.

DE Human; PRO polypeptide; secreted and transmembrane protein;

XX anti-PRO antibody; diagnostic assay; gene expression; diabetes;

KW bone disorder; cartilage disorder; rheumatoid arthritis; obesity;

KW sports injury; osteoarthritis; hyper-insulinaemia; hypo-insulinaemia;

KW hearing loss; coagulation disorder; stroke; heart attack; cardiac;

KW antidiabetic; anorectic; vulnerable; antiarthritic; osteopathic;

KW antirheumatic; auditory; cerebroprotective; angiogenic; gene; ss.

OS Homo sapiens.

XX US2003004311-A1.

PN

XX



Db 61 |||||GGCGCTGCTGGGCTGGGCTGCTCTGCTGCTGAGCTGGGCTGTGCTGGGACGAG 120  
Qy 384 GGGCCGAGGCGACCTGCGCCCCCGGGCCGACGCGCTACCACTGCTGGGAAACCTCTCGCA 443  
Db 121 GGGCCGAGGCGCACCTGCGCCCCCGGGCCGACGCGCTACCACTGCTGGGAAACCTCTCGCA 180  
Qy 444 GCTACGGCCGGGGCGCTGTATTTCAGGGCTCATCGGGCTGAGTAAGTAAGTACGACCGGT 503  
Db 181 GCTACGGCCGGGGCGCTGTATTTCAGGGCTCATCGGGCTGAGTAAGTAAGTACGACCGGT 240  
Qy 504 GTTCAACCATCTACTGGGACCTCGCGGCTGTGGTGTCTGCTGTGGGACGAGGCTGT 563  
Db 241 GTTCAACCATCTACTGGGACCTCGCGGCTGTGGTGTCTGCTGTGGGACGAGGCTGT 300  
Qy 564 GCGGGAGGCTTGGGAGGTGAGGTGAGGAGTTTCAAGCGGCGGGAAACCGTACGATGCT 623  
Db 301 GCGGGAGGCTTGGGAGGTGAGGTGAGGAGTTTCAAGCGGCGGGAAACCGTACGATGCT 360  
Qy 624 GGAAGGACCTTTGATGGCCATGGGTTTCTTCTCAACGGGAGCGGTGGAGGAGCT 683  
Db 361 GGAAGGACCTTTGATGGCCATGGGTTTCTTCTCAACGGGAGCGGTGGAGGAGCT 420  
Qy 684 GAGGAAGTTTACCATGCTTGTCTCGGGACCTGGGATGGGGAAGCGAAGGCCAGGA 743  
Db 421 GAGGAAGTTTACCATGCTTGTCTCGGGACCTGGGATGGGGAAGCGAAGGCCAGGA 480  
Qy 744 GCTGATCCAGCGGAGGCGCGGTGTCTGTGGAGACATTCAGGGGACAGAAAGGACGCC 803  
Db 481 GCTGATCCAGCGGAGGCGCGGTGTCTGTGGAGACATTCAGGGGACAGAAAGGACGCC 540  
Qy 804 ATTGATCCCTCCTGCTGTGCGCCAGGCGCACCTCCAAAGTAGTCTGCTCCCTCTT 863  
Db 541 ATTGATCCCTCCTGCTGTGCGCCAGGCGCACCTCCAAAGTAGTCTGCTCCCTCTT 600  
Qy 864 TGGCCTCCGCTTCTCTATGAGGATAGGAGTTTCCAGGCGGTGCTCCGGGACAGCTGGTG 923  
Db 601 TGGCCTCCGCTTCTCTATGAGGATAGGAGTTTCCAGGCGGTGCTCCGGGACAGCTGGTG 660  
Qy 924 TACCTGTGGGAGTACGCTCCAGGGGGGTGAGACCTAGAGATGTTCTCTGGTTCCCT 983  
Db 661 TACCTGTGGGAGTACGCTCCAGGGGGGTGAGACCTAGAGATGTTCTCTGGTTCCCT 720  
Qy 984 GCGGCCCTCGCAGGCCCCCAGAGAGCTCTCCACAGCTGAGCAGCTGGCTGCTT 1043  
Db 721 GCGGCCCTCGCAGGCCCCCAGAGAGCTCTCCACAGCTGAGCAGCTGGCTGCTT 780  
Qy 1044 CACAGTCCGGCAGGTGAGCAGCACCAAGGGGAACTTGGATGCTTGGGGCCCCGACGCTGA 1103  
Db 781 CACAGTCCGGCAGGTGAGCAGCACCAAGGGGAACTTGGATGCTTGGGGCCCCGACGCTGA 840  
Qy 1104 CTTGTGATGCTTCTGCTGAAGATGGGACAGGAGGAACAAACCCAGGACAGAAAT 1163  
Db 841 CTTGTGATGCTTCTGCTGAAGATGGGACAGGAGGAACAAACCCAGGACAGAAAT 900  
Qy 1164 CACCAACAGAAACATGCTGATGACGTATTTTCTGCTTTGCTGGGACGATGCGT 1223  
Db 901 CACCAACAGAAACATGCTGATGACGTATTTTCTGCTTTGCTGGGACGATGCGT 960  
Qy 1224 CAGCACACGGTGGGTATACCTCTGCTTCCTGATGAATAACCTCATGCTCAAAAGTG 1283  
Db 961 CAGCACACGGTGGGTATACCTCTGCTTCCTGATGAATAACCTCATGCTCAAAAGTG 1020  
Qy 1284 GGTACGTGAGAGCTGATCGGAGCTGGGGGTGGCGAGGACCAAGCTTAGGGGACCG 1343  
Db 1021 GGTACGTGAGAGCTGATCGGAGCTGGGGGTGGCGAGGACCAAGCTTAGGGGACCG 1080  
Qy 1344 TACCGGCTCCCTTACACCGACGGTTCTGATGAGGCGAGGGCTGCTGGGCTGGT 1403  
Db 1081 TACCGGCTCCCTTACACCGACGGTTCTGATGAGGCGAGGGCTGCTGGGCTGGT 1140  
Qy 1404 GCCCATGGGAATACCCCGCACCTCTATGCGGACCAACCGCTTCCGAGGCTACACCTTGC 1463

Db 1141 GCCCATGGGAATACCCCGGACCCCTCATGCGGACCAACCGGCTTCGAGGGGTACACCTGCC 1200  
Qy 1464 CCAGGGCACGGAGGTCTTCCCTCTTGGTCCATCTTGGTCCATCTGATGACCCCAACATCTTCAA 1523  
Db 1201 CCAGGGCACGGAGGTCTTCCCTCTTGGTCCATCTTGGTCCATCTGATGACCCCAACATCTTCAA 1260  
Qy 1524 GCACCCAGAGAGTTCAACCCAGACCCGTTCTCTGGATGCGATGCGGTTTCAGAGAGCA 1583  
Db 1261 GCACCCAGAGAGTTCAACCCAGACCCGTTCTCTGGATGCGATGCGGTTTCAGAGAGCA 1320  
Qy 1584 TGAGGCGTCTCTGCTCTTCTCTTAGGGAAGCGTGTCTGCTTGGAGGSCCTGGGAAA 1643  
Db 1321 TGAGGCGTCTCTGCTCTTCTCTTAGGGAAGCGTGTCTGCTTGGAGGSCCTGGGAAA 1380  
Qy 1644 AGGGAAGCTCTTCTCTTCTTCAACCACTCTCAAGCCCTTCTCTTGGAGAGCCCGGTG 1703  
Db 1381 AGGGAAGCTCTTCTCTTCTTCAACCACTCTCAAGCCCTTCTCTCTGGAGAGCCCGGTG 1440  
Qy 1704 CCGCCGGACACCTTGAGCCTCAAGCCCAAGCCCGTCAAGTGGGCTTTTCAACATTTCCCCAGC 1763  
Db 1441 CCGCCGGACACCTTGAGCCTCAAGCCCAAGCCCGTCAAGTGGGCTTTTCAACATTTCCCCAGC 1500  
Qy 1764 CTTCCAGCTGCAAGTCCGTCCTCCACTGACCTTCACTCCACAGCAGACAGATGAAGGAA 1823  
Db 1501 CTTCCAGCTGCAAGTCCGTCCTCCACTGACCTTCACTCCACAGCAGACAGATGAAGGAA 1560  
Qy 1824 GGCAACTTGAAGTGGTGGTGCCAGGACGGTGCCTTCCAGCCTCAACAGTGGGCAATGGA 1883  
Db 1561 GGCAACTTGAAGTGGTGGTGCCAGGACGGTGCCTCCAGCCTCAACAGTGGGCAATGGA 1620  
Qy 1884 CAGGTTAAGTCTCCAGAGTGTACACTGAGCAGCAGCCACATTTACAGCCTGGAGTTGT 1943  
Db 1621 CAGGTTAAGTCTCCAGAGTGTACACTGAGCAGCAGCCACATTTACAGCCTGGAGTTGT 1680  
Qy 1944 TTTCCGAGTCTCTCCACGGCCCAACGCTCACTTGACTCATGCTCTTAAGATGACAA 2003  
Db 1681 TTTCCGAGTCTCTCCACGGCCCAACGCTCACTTGACTCATGCTCTTAAGATGACAA 1740  
Qy 2004 CCGCACCCATACAACTTCAAGGGCCACAAAGCAACTGCTGGTTCAGTCTTCCACAG 2063  
Db 1741 CCGCACCCATACAACTTCAAGGGCCACAAAGCAACTGCTGGTTCAGTCTTCCACAG 1800  
Qy 2064 ACATAAATAGTTCATCTGCAATCAAGCAGCATAGCAGGTAACCCACCACTCCCT 2123  
Db 1801 ACATAAATAGTTCATCTGCAATCAAGCAGCATAGCAGGTAACCCACCACTCCCT 1860  
Qy 2124 GATCTGACGCCACAGCTGGGAGTCTGGTGTCACTTCAAGGCCACAGAAACGGCCA 2183  
Db 1861 GATCTGACGCCACAGCTGGGAGTCTGGTGTCACTTCAAGGCCACAGAAACGGCCA 1920  
Qy 2184 CACATGTTACAGCTCACAGCCCTCTTCAATTCATCGAACTTCTCAAGTGTCTTCCCT 2243  
Db 1921 CACATGTTACAGCTCACAGCCCTCTTCAATTCATCGAACTTCTCAAGTGTCTTCCCT 1980  
Qy 2244 GGTGCTGGGACAGGGAACAGCATGCCCCCTCGGGGTCAAGCCACCCAGAGACTGTCCG 2303  
Db 1981 GGTGCTGGGACAGGGAACAGCATGCCCCCTCGGGGTCAAGCCACCCAGAGACTGTCCG 2040  
Qy 2304 TGTCTATGGCCCCAACTCATGCTCTCTTGGCTACACCACTCTCCAGCCTGTGACC 2363  
Db 2041 TGTCTATGGCCCCAACTCATGCTCTCTTGGCTACACCACTCTCCAGCCTGTGACC 2100  
Qy 2364 ACCGATGTCCACACACCCCAACCACTTGTCAACAGCTTACCAGTACCGATCGTCC 2423  
Db 2101 ACCGATGTCCACACACCCCAACCACTTGTCAACAGCTTACCAGTACCGATCGTCC 2160  
Qy 2424 TGGCTCCCGAGATATTTCCCACTGAGACAGCCCGCCCCCAGAGGACAGTCCCCAG 2483  
Db 2161 TGGCTCCCGAGATATTTCCCACTGAGACAGCCCGCCCCCAGAGGACAGTCCCCAG 2220  
Qy 2484 CACCTCTGCACTGAGGCGCTCAGTACCCCTTTTTTAAAGCACCCTGATTTCTACCAATG 2543  
Db 2221 CACCTCTGCACTGAGGCGCTCAGTACCCCTTTTTTAAAGCACCCTGATTTCTACCAATG 2280

QY 2544 CAAACACATCTGGSTCTGGATTATGCACAGAGACTTTGGACATACGAGGACCCCTCAGAC 2603  
DB 2281 CAAACACATCTGGGTCGCGATTATGCACAGAGACTTTGGACATACGAGGACCCCTCAGAC 2340  
QY 2604 CGAGGAAACACCTGCGCCAAACCCCAACGCTGTATGTAAACACGTTGGAAGCGGCCCT 2663  
DB 2341 CGAGGAAACACCTGCGCCAAACCCCAACGCTGTATGTAAACACGTTGGAAGCGGCCCT 2400  
QY 2664 GTGCGCCCTCCACACACATACACACTCACTGATCTACAGCCCTGTTGCGCGTCAGAG 2723  
DB 2401 GCTGCGCCCTCCACACACATACACACTCACTGATCTACAGCCCTGTTGCGCGTCAGAG 2460  
QY 2724 TCCCACTAGACCCAGTGAAGGGTTAGAGCCAGTAGTAGGGCCAGTTTCCCAATTCACC 2783  
DB 2461 TCCCACTAGACCCAGTGAAGGGTTAGAGCCAGTAGTAGGGCCAGTTTCCCAATTCACC 2520  
QY 2784 CTGTGAGGAGTGAGCCGGATCTGACGTTCTCTTGACCTTAAGGGTCCGGCTTGGGAATT 2843  
DB 2521 CTGTGAGGAGTGAGCCGGATCTGACGTTCTCTTGACCTTAAGGGTCCGGCTTGGGAATT 2580  
QY 2844 AAAGTTTGTCTGGCCTTAGCCTAAAAAAGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 2886  
DB 2581 AAAGTTTGTCTGGCCTTAGCCTAAAAAAGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 2623

RESULT 5

ACR03634

ID ACR03634 standard; cDNA; 2623 BP.

XX ACR03634;

AC ACR03634;

DT 23-MAY-2003 (first entry)

XX cDNA encoding human PRO polypeptide #32.

XX Human; PRO polypeptide; secreted and transmembrane protein;

KW tumour necrosis factor-alpha; TNF-alpha; blood; proliferation;

KW differentiation; chondrocyte; tumour; genetic disorder; cytostatic; gene;

KW ss.

XX Homo sapiens.

XX US2003036180-A1.

XX 20-FEB-2003.

XX 09-MAY-2002; 2002US-00143114.

XX 31-MAR-1997; 97WO-US005230.

PR 12-JUN-1998; 98WO-US012456.

PR 14-JUL-1998; 98WO-US014552.

PR 28-AUG-1998; 98WO-US017888.

PR 10-SEP-1998; 98WO-US018824.

PR 14-SEP-1998; 98WO-US019094.

PR 14-SEP-1998; 98WO-US019177.

PR 16-SEP-1998; 98WO-US019330.

PR 17-SEP-1998; 98WO-US019437.

PR 07-OCT-1998; 98WO-US021141.

PR 29-OCT-1998; 98WO-US022991.

PR 20-NOV-1998; 98WO-US024855.

PR 01-DEC-1998; 98WO-US025108.

PR 05-JAN-1999; 99WO-US000106.

PR 08-MAR-1999; 99WO-US005028.

PR 10-MAR-1999; 99WO-US005190.

PR 20-APR-1999; 99WO-US008615.

PR 14-MAY-1999; 99WO-US010733.

PR 02-JUN-1999; 99WO-US012252.

PR 01-SEP-1999; 99WO-US020111.

PR 08-SEP-1999; 99WO-US020594.

PR 13-SEP-1999; 99WO-US020944.

PR 15-SEP-1999; 99WO-US021090.

PR 15-SEP-1999; 99WO-US021547.

PR 05-OCT-1999; 99WO-US023089.

PR 29-NOV-1999; 99WO-US028214.

PR 30-NOV-1999; 99WO-US028313.

PR 30-NOV-1999; 99WO-US028409.

PR 01-DEC-1999; 99WO-US028301.

PR 01-DEC-1999; 99WO-US028634.

PR 02-DEC-1999; 99WO-US028551.

PR 02-DEC-1999; 99WO-US028564.

PR 02-DEC-1999; 99WO-US028565.

PR 16-DEC-1999; 99WO-US030095.

PR 20-DEC-1999; 99WO-US030911.

PR 20-DEC-1999; 99WO-US030999.

PR 22-DEC-1999; 99WO-US030720.

PR 30-DEC-1999; 99WO-US031243.

PR 30-DEC-1999; 99WO-US031274.

PR 05-JAN-2000; 2000WO-US000219.

PR 06-JAN-2000; 2000WO-US000277.

PR 06-JAN-2000; 2000WO-US000376.

PR 11-FEB-2000; 2000WO-US000365.

PR 11-FEB-2000; 2000WO-US004341.

PR 18-FEB-2000; 2000WO-US004342.

PR 22-FEB-2000; 2000WO-US004414.

PR 24-FEB-2000; 2000WO-US004914.

PR 24-FEB-2000; 2000WO-US005004.

PR 01-MAR-2000; 2000WO-US005601.

PR 02-MAR-2000; 2000WO-US005746.

PR 02-MAR-2000; 2000WO-US005841.

PR 10-MAR-2000; 2000WO-US006319.

PR 15-MAR-2000; 2000WO-US006884.

PR 20-MAR-2000; 2000WO-US007377.

PR 21-MAR-2000; 2000WO-US007532.

PR 30-MAR-2000; 2000WO-US008439.

PR 17-MAY-2000; 2000WO-US013705.

PR 22-MAY-2000; 2000WO-US014042.

PR 30-MAY-2000; 2000WO-US014941.

PR 02-JUN-2000; 2000WO-US015264.

PR 28-JUL-2000; 2000WO-US020710.

PR 11-AUG-2000; 2000WO-US022031.

PR 23-AUG-2000; 2000WO-US023522.

PR 24-AUG-2000; 2000WO-US023328.

PR 08-NOV-2000; 2000WO-US030952.

PR 10-NOV-2000; 2000WO-US030873.

PR 01-DEC-2000; 2000WO-US032678.

PR 20-DEC-2000; 2000US-00747259.

PR 20-DEC-2000; 2000WO-US034956.

PR 28-FEB-2001; 2001US-00796498.

PR 28-FEB-2001; 2001WO-US006520.

PR 01-MAR-2001; 2001WO-US006666.

PR 09-MAR-2001; 2001US-00802706.

PR 14-MAR-2001; 2001US-00808689.

PR 22-MAR-2001; 2001US-00816744.

PR 05-APR-2001; 2001US-00828366.

PR 10-MAY-2001; 2001US-00854208.

PR 10-MAY-2001; 2001US-00854280.

PR 18-MAY-2001; 2001US-00860216.

PR 25-MAY-2001; 2001US-00866028.

PR 25-MAY-2001; 2001US-00866034.

PR 25-MAY-2001; 2001WO-US017092.

PR 01-JUN-2001; 2001US-00872035.

PR 01-JUN-2001; 2001WO-US017800.

PR 05-JUN-2001; 2001US-00874503.

PR 14-JUN-2001; 2001US-00882636.

PR 19-JUN-2001; 2001US-00886342.

PR 20-JUN-2001; 2001WO-US019692.

PR 21-JUN-2001; 2001US-00887879.

PR 22-JUN-2001; 2001WO-US020116.

PR 29-JUN-2001; 2001WO-US021066.

PR 09-JUL-2001; 2001WO-US021735.

PR 18-JUL-2001; 2001US-00908827.

PR 06-AUG-2001; 2001US-00924419.

PR 09-AUG-2001; 2001US-00927796.









Db 541 ATTCGATCCCTCCCTGCTGCTGGCCGACGACCCTCCAAACGCTAGTCTGTGCTCCCTCCTCTT 600  
Qy 864 TGGCCCTCGCTTCTCTATGAGGATAAGAGTTTCAAGGCCGTGTCTCCGGGACAGCTGGTG 923  
Db 601 TGGCCCTCGCTTCTCTATGAGGATAAGAGTTTCAAGGCCGTGTCTCCGGGACAGCTGGTG 660  
Qy 924 TACCCCTGCTGGAGTCAAGTCCAGGGGGGTGAGACCTACGAGATGTTCTCTCGGTTCT 983  
Db 661 TACCCCTGCTGGAGTCAAGTCCAGGGGGGTGAGACCTACGAGATGTTCTCTCGGTTCT 720  
Qy 984 GGGGCCCTGCGAGGCCCCCAAGCAGTCTCTCCACACGTCAGCACCTTGGCTGCCTT 1043  
Db 721 GGGGCCCTGCGAGGCCCCCAAGCAGTCTCTCCACACGTCAGCACCTTGGCTGCCTT 780  
Qy 1044 CACAGTCCGGCAGGTGCAGCAGCACAGGGGAACTGTGATGCTTGGGCCCGCGCAGTGA 1103  
Db 781 CACAGTCCGGCAGGTGCAGCAGCACAGGGGAACTGTGATGCTTGGGCCCGCGCAGTGA 840  
Qy 1104 CTTTGTGATGCTTCTCTGCTGAAGATGGCACAGGAGGAACAAAACCCAGGCACAGAAATT 1163  
Db 841 CTTTGTGATGCTTCTCTGCTGAAGATGGCACAGGAGGAACAAAACCCAGGCACAGAAATT 900  
Qy 1164 CACCAACAGAACATGCTGATGACAGTCAATTTATTTGCTGTTGTGGACGATGACGGT 1223  
Db 901 CACCAACAGAACATGCTGATGACAGTCAATTTATTTGCTGTTGTGGACGATGACGGT 960  
Qy 1224 CAGCACCGGTGGGTATACCTCTGCTGCTGATGAATACCTCATGTCACAAAGTG 1283  
Db 961 CAGCACCGGTGGGTATACCTCTGCTGCTGATGAATACCTCATGTCACAAAGTG 1020  
Qy 1284 GGTACGTGAGGAGCTGAATCGGGAGCTGGGGGCTGGCCAGGCACCAAGCTTAGGGGACCG 1343  
Db 1021 GGTACGTGAGGAGCTGAATCGGGAGCTGGGGGCTGGCCAGGCACCAAGCTTAGGGGACCG 1080  
Qy 1344 TACCCGCTCCCTTACACGACGCGTTCGTGATGAGGGGACGCGCTGCTGGCGTGGT 1403  
Db 1081 TACCCGCTCCCTTACACGACGCGTTCGTGATGAGGGGACGCGCTGCTGGCGTGGT 1140  
Qy 1404 GCCCATGGGAATACCCGACACCTCATGCGGACCAACCGCTTCGAGGGTACACCTGCC 1463  
Db 1141 GCCCATGGGAATACCCGACACCTCATGCGGACCAACCGCTTCGAGGGTACACCTGCC 1200  
Qy 1464 CAGGGCACGAGGTCTTCCCTCTCTGCTGCTCCATCTCTGATGACCCCAACATCTTCAA 1523  
Db 1201 CCAGGGCACGAGGTCTTCCCTCTCTGCTGCTCCATCTCTGATGACCCCAACATCTTCAA 1260  
Qy 1524 GCACCCAGAGATTCAACCCACAGCGTTTCTGATGAGATGACGCGTTTCAAGGAACA 1583  
Db 1261 GCACCCAGAGATTCAACCCACAGCGTTTCTGATGAGATGACGCGTTTCAAGGAACA 1320  
Qy 1584 TGAGGCGTTCCCTGCGCTTCTCTTAGGGAAGGTGTCTGCTTGGAGAGGGCTTGGCAAA 1643  
Db 1321 TGAGGCGTTCCCTGCGCTTCTCTTAGGGAAGGTGTCTGCTTGGAGAGGGCTTGGCAAA 1380  
Qy 1644 AGCGAGTCTTCTCTTCTTCAACCACTCTTCAAGCCCTTCTCTGAGAGCCCGTG 1703  
Db 1381 AGCGAGTCTTCTCTTCTTCAACCACTCTTCAAGCCCTTCTCTGAGAGCCCGTG 1440  
Qy 1704 CCCGGGACACCTGAGCTCAAGCCACCGTCTGATGGCTTTTCAACATTTCCCGCAGC 1763  
Db 1441 CCCGGGACACCTGAGCTCAAGCCACCGTCTGATGGCTTTTCAACATTTCCCGCAGC 1500  
Qy 1764 CTTCCAGTGTCAAGTCCGCTCCCTGACCTTCACTCCACACGACAGACAGATGAAGAA 1823  
Db 1501 CTTCCAGTGTCAAGTCCGCTCCCTGACCTTCACTCCACACGACAGACAGATGAAGAA 1560  
Qy 1824 GCGAATTGGAAGTGGTGGTCCCGACGACGCTGCTCCAGCTTCAACAGTGGGATGGA 1883  
Db 1561 GCGAATTGGAAGTGGTGGTCCCGACGACGCTGCTCCAGCTTCAACAGTGGGATGGA 1620  
Qy 1884 CAGGTTAATGCTCCAGTGTACACTGACGGGACGACATTTACAGCCTGCAAGTTGT 1943  
Db 1621 CAGGTTAATGCTCCAGAGTGTACACTGACGGGACGACATTTACAGCCTGCAAGTTGT 1680

Qy 1944 TTTTCGGAGTCTGTCCACGGCCCAACGCTCACTTGACTCATGCTTAAGATGCACAA 2003  
Db 1681 TTTTCGGAGTCTGTCCACGGCCCAACGCTCACTTGACTCATGCTTAAGATGCACAA 1740  
Qy 2004 CCGCACACCCATACACAACTACAAGGGCCCAAAAGCAACTGTGGTTAGCTTTTCACAG 2063  
Db 1741 CCGCACACCCATACACAACTACAAGGGCCCAAAAGCAACTGTGGTTAGCTTTTCACAG 1800  
Qy 2064 ACATAAATATAGTCCATCTGCAATCAAAAGCAATAGCCAGGTAAACCAACTCCCT 2123  
Db 1801 ACATAAATATAGTCCATCTGCAATCAAAAGCAATAGCCAGGTAAACCAACTCCCT 1860  
Qy 2124 GGATCTGAGCCCAACGTTGGGAGTCTGGCTGTCACTTCAAGCCACAGAAAGGCCA 2183  
Db 1861 GGATCTGAGCCCAACGTTGGGAGTCTGGCTGTCACTTCAAGCCACAGAAAGGCCA 1920  
Qy 2184 CACATGTTTACAGCTCACAGCCCTCTCCATTTATCGAACTTCTCAGTGTCCCTGTCCT 2243  
Db 1921 CACATGTTTACAGCTCACAGCCCTCTCCATTTATCGAACTTCTCAGTGTCCCTGTCCT 1980  
Qy 2244 GGTGCTTGGCACAGGGAACAGCATGCCCTCTCCGGGTCATGCCACCCAGAGCTGTCC 2303  
Db 1981 GGTGCTTGGCACAGGGAACAGCATGCCCTCTCCGGGTCATGCCACCCAGAGCTGTCC 2040  
Qy 2304 TGTCTATGGCCCCAACTCATGCTCCCTCTTGGCTACACACTCTCCAGCTGTGACC 2363  
Db 2041 TGTCTATGGCCCCAACTCATGCTCCCTCTTGGCTACACACTCTCCAGCTGTGACC 2100  
Qy 2364 AGCGATGTCCACACACCCCAACCACTTGTCCACAGCTACCCAGCTACGACATGTC 2423  
Db 2101 AGCGATGTCCACACACCCCAACCACTTGTCCACAGCTACCCAGCTACGACATGTC 2160  
Qy 2424 TGGCTCCCGAGATATTTCCACTGAGACAGCGCCGCCCAACAGAGGACAGTCCCGAG 2483  
Db 2161 TGGCTCCCGAGATATTTCCACTGAGACAGCGCCGCCCAACAGAGGACAGTCCCGAG 2220  
Qy 2484 CCACCTCTGCAACTGACGCCCTCAGTCAACCCCTTTTAAAGCACCTGATTTACCAATG 2543  
Db 2221 CCACCTCTGCAACTGACGCCCTCAGTCAACCCCTTTTAAAGCACCTGATTTACCAATG 2280  
Qy 2544 CAAACACATCTGGGTCTCGATTTATGACAGAGACTTTGGACATACGAGGACCTCAGAC 2603  
Db 2281 CAAACACATCTGGGTCTCGATTTATGACAGAGACTTTGGACATACGAGGACCTCAGAC 2340  
Qy 2604 CGGAGAACACTGCCCCAACCCCAACAGTGTATGTAAACACGTGGAAAGCGGCCCT 2663  
Db 2341 CGGAGGAACAACCTGCCCAACCCCAACAGTGTATGTAAACCAACGTGGAAAGCGGCCCT 2400  
Qy 2664 GCTGCCCTTCCACACACACATACACTCACTGATCTACAGCCCTGTTCCGCGTCAGAG 2723  
Db 2401 GCTGCCCTTCCACACACACATACACTCACTGATCTACAGCCCTGTTCCGCGTCAGAG 2460  
Qy 2724 TCCCACTAGACCCAGTGGAAAGGGTTAGAGACCAAGTAGGGGCGAGTTTCAAATTCAC 2783  
Db 2461 TCCCACTAGACCCAGTGGAAAGGGTTAGAGACCAAGTAGGGGCGAGTTTCAAATTCAC 2520  
Qy 2784 CTGTTCAGGGAGTGAAGCGGATCTGAGCTTCTTGTGACTTAAGGGTCCGGCTTGGGAAT 2843  
Db 2521 CTGTTCAGGGAGTGAAGCGGATCTGAGCTTCTTGTGACTTAAGGGTCCGGCTTGGGAAT 2580  
Qy 2844 AAAAGTTTCTTGTGCGCTTTTAGCCTTAAAAAATAAAAAAAAAAAAA 2886  
Db 2581 AAAAGTTTCTTGTGCGCTTTAGCCTTAAAAAATAAAAAAAAAAAAA 2623

## RESULT 7

ACD41826

ID ACD41826 standard; cDNA; 2623 BP.

XX ACD41826;

XX ACD41826;

DT 05-SEP-2003 (first entry)

XX DE Human secreted/transmembrane protein (PRO) cDNA #32.  
XX KW Human, ss; gene; PRO; secreted protein; transmembrane protein; tumour;  
KW proteolytic; gene therapy; tumour necrosis factor-alpha; TNF-alpha; blood;  
KW chondrocyte; cartilage; cytokine; peripheral blood mononuclear cell;  
KW BMC; glucose uptake; FFA; skeletal muscle cell; adipocyte cell;  
KW chondrocyte cell proliferation; chondrocyte cell differentiation;  
KW pericyte cell; inner ear utricular supporting cell; r-lymphocyte cell;  
KW endothelial cell; A-peptide; factor VIIA.  
XX OS Homo sapiens.  
XX PN US2003036179-A1.  
XX PD 20-FEB-2003.  
XX PF 10-MAY-2002; 2002US-00142431.  
XX PR 31-MAR-1997; 97WO-US005230.  
PR 12-JUN-1998; 98WO-US012456.  
PR 14-JUL-1998; 98WO-US014552.  
PR 28-AUG-1998; 98WO-US017888.  
PR 10-SEP-1998; 98WO-US018624.  
PR 14-SEP-1998; 98WO-US019093.  
PR 14-SEP-1998; 98WO-US019094.  
PR 14-SEP-1998; 98WO-US019177.  
PR 16-SEP-1998; 98WO-US019330.  
PR 17-SEP-1998; 98WO-US019437.  
PR 07-OCT-1998; 98WO-US021141.  
PR 29-OCT-1998; 98WO-US022991.  
PR 29-OCT-1998; 98WO-US022992.  
PR 20-NOV-1998; 98WO-US024855.  
PR 01-DEC-1998; 98WO-US025108.  
PR 05-JAN-1999; 99WO-US000106.  
PR 08-MAR-1999; 99WO-US005028.  
PR 10-MAR-1999; 99WO-US005190.  
PR 20-APR-1999; 99WO-US008615.  
PR 14-MAY-1999; 99WO-US010733.  
PR 02-JUN-1999; 99WO-US012252.  
PR 08-SEP-1999; 99WO-US020111.  
PR 08-SEP-1999; 99WO-US020594.  
PR 13-SEP-1999; 99WO-US021090.  
PR 15-SEP-1999; 99WO-US021094.  
PR 15-SEP-1999; 99WO-US021547.  
PR 05-OCT-1999; 99WO-US023089.  
PR 29-NOV-1999; 99WO-US028214.  
PR 30-NOV-1999; 99WO-US028313.  
PR 30-NOV-1999; 99WO-US028409.  
PR 01-DEC-1999; 99WO-US028301.  
PR 01-DEC-1999; 99WO-US028634.  
PR 02-DEC-1999; 99WO-US028551.  
PR 02-DEC-1999; 99WO-US028564.  
PR 02-DEC-1999; 99WO-US028565.  
PR 16-DEC-1999; 99WO-US030095.  
PR 20-DEC-1999; 99WO-US030391.  
PR 20-DEC-1999; 99WO-US030599.  
PR 22-DEC-1999; 99WO-US030720.  
PR 30-DEC-1999; 99WO-US031243.  
PR 30-DEC-1999; 99WO-US031274.  
PR 05-JAN-2000; 2000WO-US000219.  
PR 06-JAN-2000; 2000WO-US000277.  
PR 06-JAN-2000; 2000WO-US000376.  
PR 11-FEB-2000; 2000WO-US003565.  
PR 18-FEB-2000; 2000WO-US004341.  
PR 18-FEB-2000; 2000WO-US004342.  
PR 22-FEB-2000; 2000WO-US004414.  
PR 24-FEB-2000; 2000WO-US004914.  
PR 24-FEB-2000; 2000WO-US005004.  
PR 01-MAR-2000; 2000WO-US005601.  
PR 02-MAR-2000; 2000WO-US005746.  
PR 02-MAR-2000; 2000WO-US005841.  
PR 10-MAR-2000; 2000WO-US006319.  
PR 15-MAR-2000; 2000WO-US006884.  
PR 20-MAR-2000; 2000WO-US007377.  
PR 21-MAR-2000; 2000WO-US007532.  
PR 30-MAR-2000; 2000WO-US008439.  
PR 17-MAY-2000; 2000WO-US013705.  
PR 22-MAY-2000; 2000WO-US014042.  
PR 30-MAY-2000; 2000WO-US014941.  
PR 02-JUN-2000; 2000WO-US015264.  
PR 28-JUN-2000; 2000WO-US020710.  
PR 11-AUG-2000; 2000WO-US022031.  
PR 23-AUG-2000; 2000WO-US023522.  
PR 24-AUG-2000; 2000WO-US023328.  
PR 08-NOV-2000; 2000WO-US030352.  
PR 10-NOV-2000; 2000WO-US030873.  
PR 01-DEC-2000; 2000WO-US032678.  
PR 20-DEC-2000; 2000US-00747259.  
PR 20-DEC-2000; 2000WO-US034956.  
PR 28-FEB-2001; 2001US-00796498.  
PR 28-FEB-2001; 2001US-00796498.  
PR 01-MAR-2001; 2001WO-US006520.  
PR 01-MAR-2001; 2001WO-US006666.  
PR 14-MAR-2001; 2001US-00802706.  
PR 22-MAR-2001; 2001US-00808689.  
PR 05-APR-2001; 2001US-00816744.  
PR 10-MAY-2001; 2001US-00828166.  
PR 10-MAY-2001; 2001US-00854208.  
PR 18-MAY-2001; 2001US-00854280.  
PR 18-MAY-2001; 2001US-00860216.  
PR 25-MAY-2001; 2001US-00866034.  
PR 25-MAY-2001; 2001US-00866034.  
PR 01-JUN-2001; 2001US-00872035.  
PR 01-JUN-2001; 2001WO-US017800.  
PR 05-JUN-2001; 2001US-00874503.  
PR 14-JUN-2001; 2001US-00882636.  
PR 19-JUN-2001; 2001US-00886342.  
PR 20-JUN-2001; 2001WO-US019692.  
PR 21-JUN-2001; 2001US-00887879.  
PR 22-JUN-2001; 2001WO-US020116.  
PR 29-JUN-2001; 2001WO-US021066.  
PR 09-JUL-2001; 2001WO-US021735.  
PR 18-JUL-2001; 2001US-00908827.  
PR 06-AUG-2001; 2001US-00924419.  
PR 09-AUG-2001; 2001US-00927796.  
PR 16-AUG-2001; 2001US-00931836.  
PR 19-DEC-2001; 2001US-00028072.  
XX XX (GETH ) GENENTECH INC.  
PA Baker KP, Beresini M, Deforge L, Desnoyers L, Filvaroff E, Gao W;  
PI Gerritsen ME, Goddard A, Godowski PU, Gurney AL, Sherwood S;  
PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;  
XX WPI; 2003-466355/44.  
DR P-PSDB; ABO24872.  
DR PT New isolated nucleic acid encoding a PRO polypeptide, e.g. PRO1114 or  
PT PRO4978, useful in molecular biology, chromosome and gene mapping, in  
PT generating antisense RNA and DNA, and in gene therapy.  
XX Claim 2; Fig 63; 659pp; English.  
XX The invention relates to an isolated nucleic acid comprising at least 80%  
CC sequence identity to a PRO (secreted and transmembrane protein) cDNA  
CC comprising a nucleic acid (a) encoding a PRO polypeptide, or its  
CC extracellular domain (with or without its associated signal peptide),  
CC which comprises any of the 275 120-850 residue amino acid sequences,  
CC given in the specification; (b) comprising any of the 275 300-3500  
CC nucleotide sequences, given in the specification; or (c) comprising the  
CC full-length coding sequence of the nucleotide sequences given in the  
CC specification, or of the DNA deposited under any of the American Type  
CC Culture Collection (ATCC) Accession Numbers listed in the specification.  
CC Also included are a vector comprising the novel nucleic acid, a host cell  
CC comprising the vector, producing a PRO polypeptide, the isolated PRO



```
Db      1681  TTTCCGGAGTCTGTCCTCCAGGCCACACAGCTCACTTGACTATGCTGCTAAGTGCACAA 1740
Qy      2004  CCGCACACCCATACACAACTACAAGGGCCACAAGCAAACTGCTGGGTTAGCTTTCCACAG 2063
Db      1741  CCGCACACCCATACACAACTACAAGGGCCACAAGCAAACTGCTGGGTTAGCTTTCCACAG 1800
Qy      2064  ACATAAATATAGTCCATCTGCAATCAAGACATAGCCAGGTAAACCCACCACTCCCT 2123
Db      1801  ACATAAATATAGTCCATCTGCAATCAAGACATAGCCAGGTAAACCCACCACTCCCT 1860
Qy      2124  GGATCTGCAGCCACACAGTGGGAGTCTGGCTGTCACTTCAAGCCACAGAAACGGCCA 2183
Db      1861  GGATCTGCAGCCACACAGTGGGAGTCTGGCTGTCACTTCAAGCCACAGAAACGGCCA 1920
Qy      2184  CACATGTTCAAGCTCACAGCCCTCTCCATCTCATTCGAATCTTCAGTGTCCCTGTCCCT 2243
Db      1921  CACATGTTCAAGCTCACAGCCCTCTCCATCTCATTCGAATCTTCAGTGTCCCTGTCCCT 1980
Qy      2244  GGTGCTGGCACAGGACAGCATGCCCTCCGGGTCATGCCACCCAGAGACTGTGCG 2303
Db      1981  GGTGCTGGCACAGGAAACAGCATGCCCTCCGGGTCATGCCACCCAGAGACTGTGCG 2040
Qy      2304  TGTCTATGGCCCCCACTCATGCTCCCTCTCTTGGCTACACCACTCTCCAGCCTGTGACC 2363
Db      2041  TGTCTATGGCCCCCACTCATGCTCCCTCTCTTGGCTACACCACTCTCCAGCCTGTGACC 2100
Qy      2364  ACCGATGTCACACACCCCAACCACTTGTGTCACACAGCTACCCAGTACGACATCGTCC 2423
Db      2101  ACCGATGTCACACACCCCAACCACTTGTGTCACACAGCTACCCAGTACGACATCGTCC 2160
Qy      2424  TGGCTCCCCAGAGTATCTTCCACTGAGACAGCGGCCCCACAGAGGCACAGTCCCCAG 2483
Db      2161  TGGCTCCCCAGAGTATCTTCCACTGAGACAGCGGCCCCACAGAGGCACAGTCCCCAG 2220
Qy      2484  CCACCTCTGCAACTGCAGCCCTCAGTCAACCCCTTTTAAAGCACCTTGATTTACCAATG 2543
Db      2221  CCACCTCTGCAACTGCAGCCCTCAGTCAACCCCTTTTAAAGCACCTTGATTTACCAATG 2280
Qy      2544  CAACACATCTGGTCTGGGATTATGCACAGAGCTTTGGACATACGAGGACCTTCAGAC 2603
Db      2281  CAACACATCTGGTCTGGGATTATGCACAGAGCTTTGGACATACGAGGACCTTCAGAC 2340
Qy      2604  CGGAGGAACACTCTCCCAACCCCAACAGCTGCTTATGTAAACAGTGGAAAGCGGCCCT 2663
Db      2341  CGGAGGAACACTCTCCCAACCCCAACAGCTGCTTATGTAAACAGTGGAAAGCGGCCCT 2400
Qy      2664  GCTGCCCTCCACACACATACACACTCACTGATCTACAGCCCTGTTTCGGCGTCAAG 2723
Db      2401  GCTGCCCTCCACACACATACACACTCACTGATCTACAGCCCTGTTTCGGCGTCAAG 2460
Qy      2724  TCCCACTAGACCCAGTGGAGGGTTAGACACCAAGTAGGGGCAGTTTCCAAATTCACC 2783
Db      2461  TCCCACTAGACCCAGTGGAGGGTTAGACACCAAGTAGGGGCAGTTTCCAAATTCACC 2520
Qy      2784  CTGTCAGGAGTGAAGCGGATCTCAGCTTCTGCTGATTAAGGTCGGCTTCGGGAAAT 2843
Db      2521  CTGTCAGGAGTGAAGCGGATCTCAGCTTCTGCTGATTAAGGTCGGCTTCGGGAAAT 2580
Qy      2844  AAAGTTTGTTCGGCTTTAGCCTAAAAAATAAAAAAAAAA 2886
Db      2581  AAAGTTTGTTCGGCTTTAGCCTAAAAAATAAAAAAAAAA 2623
```

## RESULT 8

ACA04055

ID ACA04055 standard; cDNA; 2623 BP.

XX ACA04055;

AC ACA04055;

XX 27-MAY-2003 (first entry)

DT Human cDNA encoding a secreted/transmembrane protein, SEQ ID 63.

XX

DE

XX Human; ss; gene; secreted protein; transmembrane protein; PRO; inflammatory disease; organ failure; atherosclerosis; cardiac injury; infertility; birth defects; premature aging; AIDS; biosensor; acquired immunodeficiency syndrome; cancer; diabetic complication; bioreactor; tumour.

OS Homo sapiens.

XX US2003032155-A1.

XX 13-FEB-2003.

PF 03-MAY-2002; 2002US-00137865.

XX 31-MAR-1997; 97WO-US005230.

PR 12-JUN-1998; 98WO-US012456.

PR 14-JUL-1998; 98WO-US014552.

PR 28-AUG-1998; 98WO-US017888.

PR 10-SEP-1998; 98WO-US018824.

PR 14-SEP-1998; 98WO-US019093.

PR 14-SEP-1998; 98WO-US019094.

PR 14-SEP-1998; 98WO-US019177.

PR 16-SEP-1998; 98WO-US019330.

PR 17-SEP-1998; 98WO-US019437.

PR 07-OCT-1998; 98WO-US021141.

PR 29-OCT-1998; 98WO-US022991.

PR 29-OCT-1998; 98WO-US022992.

PR 20-NOV-1998; 98WO-US024855.

PR 01-DEC-1998; 98WO-US025108.

PR 05-JAN-1999; 99WO-US000106.

PR 08-MAR-1999; 99WO-US005028.

PR 10-MAR-1999; 99WO-US005190.

PR 20-APR-1999; 99WO-US008615.

PR 14-MAY-1999; 99WO-US010733.

PR 02-JUN-1999; 99WO-US012252.

PR 01-SEP-1999; 99WO-US020111.

PR 08-SEP-1999; 99WO-US020594.

PR 13-SEP-1999; 99WO-US020944.

PR 15-SEP-1999; 99WO-US021090.

PR 05-OCT-1999; 99WO-US021547.

PR 29-NOV-1999; 99WO-US028214.

PR 30-NOV-1999; 99WO-US028313.

PR 30-NOV-1999; 99WO-US028409.

PR 01-DEC-1999; 99WO-US028301.

PR 02-DEC-1999; 99WO-US028551.

PR 02-DEC-1999; 99WO-US028564.

PR 02-DEC-1999; 99WO-US028565.

PR 16-DEC-1999; 99WO-US030095.

PR 20-DEC-1999; 99WO-US030911.

PR 20-DEC-1999; 99WO-US030999.

PR 22-DEC-1999; 99WO-US030720.

PR 30-DEC-1999; 99WO-US031243.

PR 30-DEC-1999; 99WO-US031274.

PR 05-JAN-2000; 2000WO-US000219.

PR 06-JAN-2000; 2000WO-US000277.

PR 06-JAN-2000; 2000WO-US000376.

PR 11-FEB-2000; 2000WO-US003565.

PR 18-FEB-2000; 2000WO-US004341.

PR 18-FEB-2000; 2000WO-US004342.

PR 22-FEB-2000; 2000WO-US004414.

PR 24-FEB-2000; 2000WO-US004914.

PR 24-FEB-2000; 2000WO-US005004.

PR 01-MAR-2000; 2000WO-US005601.

PR 02-MAR-2000; 2000WO-US005746.

PR 02-MAR-2000; 2000WO-US005841.

PR 10-MAR-2000; 2000WO-US006319.

PR 15-MAR-2000; 2000WO-US006884.

PR 20-MAR-2000; 2000WO-US007377.

PR 21-MAR-2000; 2000WO-US007532.

PR 30-MAR-2000; 2000WO-US008439.

PR 17-MAY-2000; 2000WO-US013705.  
PR 22-MAY-2000; 2000WO-US014042.  
PR 30-MAY-2000; 2000WO-US014941.  
PR 02-JUN-2000; 2000WO-US015264.  
PR 28-JUL-2000; 2000WO-US020710.  
PR 11-AUG-2000; 2000WO-US022031.  
PR 23-AUG-2000; 2000WO-US023522.  
PR 24-AUG-2000; 2000WO-US023328.  
PR 08-NOV-2000; 2000WO-US030952.  
PR 10-NOV-2000; 2000WO-US030873.  
PR 01-DEC-2000; 2000WO-US032678.  
PR 20-DEC-2000; 2000US-00747259.  
PR 20-DEC-2000; 2000WO-US034956.  
PR 28-FEB-2001; 2000US-00796498.  
PR 28-FEB-2001; 2001WO-US006520.  
PR 01-MAR-2001; 2001WO-US006666.  
PR 09-MAR-2001; 2001US-00802706.  
PR 14-MAR-2001; 2001US-00808689.  
PR 22-MAR-2001; 2001US-00815744.  
PR 05-APR-2001; 2001US-00828366.  
PR 10-MAY-2001; 2001US-00854208.  
PR 18-MAY-2001; 2001US-00860216.  
PR 25-MAY-2001; 2001US-00866028.  
PR 25-MAY-2001; 2001US-00866034.  
PR 25-MAY-2001; 2001WO-US017092.  
PR 01-JUN-2001; 2001US-00872035.  
PR 01-JUN-2001; 2001WO-US017800.  
PR 05-JUN-2001; 2001US-00874503.  
PR 14-JUN-2001; 2001US-00882636.  
PR 19-JUN-2001; 2001US-00886342.  
PR 20-JUN-2001; 2001WO-US019692.  
PR 21-JUN-2001; 2001US-00887879.  
PR 22-JUN-2001; 2001WO-US020116.  
PR 29-JUN-2001; 2001WO-US021066.  
PR 09-JUL-2001; 2001WO-US021735.  
PR 18-JUL-2001; 2001US-00908827.  
PR 06-AUG-2001; 2001US-00924419.  
PR 09-AUG-2001; 2001US-00927796.  
PR 16-AUG-2001; 2001US-00931836.  
PR 19-DEC-2001; 2001US-00028072.

(GETH ) GENENTECH INC.

XX Baker KP, Beresini M, Deforge L, Desnoyers L, Filvaroff E, Cao W;  
PI Gerritsen MB, Goddard A, Godowski PJ, Gurney AL, Sherwood S;  
PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;  
XX

DR WPI; 2003-331925/31.

DR P-PSDB; ABU66877.

XX New secreted and transmembrane nucleic acids and polypeptides, designated  
PT as PRO, useful for treating inflammation, organ failure, atherosclerosis,  
PT cardiac injury, infertility, birth defects, premature aging, AIDS, or  
PT cancer.

XX Claim 2; Fig 63; 659pp; English.

XX The invention relates to an isolated nucleic acid comprising, or which is  
CC at least 80% identical to, or the full-length coding sequence of, any of  
CC the 275 nucleotide sequences, encoding the corresponding PRO polypeptide  
CC (one of 275 secreted or transmembrane proteins). The nucleic acid further  
CC comprises the full-length coding sequence of the DNA deposited under  
CC American Type Culture Collection (ATCC) accession number in a list given  
CC in the specification. Also included are vectors and host cells for  
CC producing PRO proteins, PRO fusion proteins, anti-PRO antibodies, PRO  
CC extracellular domains and mature sequences, methods of detecting PRO  
CC proteins, methods for stimulating the release of TNF-alpha (tumour  
CC necrosis factor alpha) from human blood, (and the proliferation of  
CC differentiation of chondrocyte cells, the proliferation of, or gene  
CC expression in pericyte cells, the release of proteoglycans from  
CC cartilage, proliferation of inner ear articular supporting cells, the  
CC proliferation of T-lymphocyte cells, the release of a cytokine from

CC peripheral blood mononuclear cells (PBMC), or the proliferation of  
CC endothelial cells), a method for modulating the uptake of glucose or free  
CC fatty acid (FFA) by skeletal muscle cells, a method for inhibiting the  
CC binding of A-peptide to factor VIIA, or the differentiation of adipocyte  
CC cells, a method for detecting the presence of a tumour in a mammal and an  
CC oligonucleotide probe derived from any of the nucleotide sequences cited  
CC above. The nucleic acids and polypeptides are useful for treating  
CC inflammatory diseases, organ failure, atherosclerosis, cardiac injury,  
CC infertility, birth defects, premature aging, AIDS (acquired  
CC immunodeficiency syndrome), cancer, or diabetic complications. The  
CC nucleic acids are useful as hybridisation probes, in chromosome and gene  
CC mapping, and in generating antisense RNA or DNA. The polypeptides are  
CC useful as pharmaceuticals, diagnostics, biosensors or bioreactors. Both  
CC are useful in tissue typing. The present sequence encodes a PRO protein  
CC of the invention

XX Sequence 2623 BP; 552 A; 865 C; 704 G; 502 T; 0 U; 0 Other;

Query Match 89.0%; Score 2621.4; DB 7; Length 2623;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2622; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 264 GCGAGCGCTGGAGAGAGAGAGCGACCTGCCAGATGAGCGGACCGCACCTG 323  
DB 1 GCGGAGCGCTGGAGAGAGAGAGCGACCTGCCAGATGAGCGGACCGCACCTG 60  
QY 324 GCGGCTGCTGGCGCTGGCGCTGCTCTGCTGTGACGCTGGCGCTGTCGGGACCG 383  
DB 61 GCGGCTGCTGGCGCTGGCGCTGCTCTGCTGTGACGCTGGCGCTGTCGGGACCG 120  
QY 384 GCGCGGAGGCGACCTGCCCGCCCGGCGCCACGCGCTTACCACCTGCTGGGAAACCTCTCGCA 443  
DB 121 GCGCGGAGGCGACCTGCCCGCCCGGCGCCACGCGCTTACCACCTGCTGGGAAACCTCTCGCA 180  
QY 444 GCTACGCGCGGCGCGCTCTATTACGGGCTCATATGCGGCTGAGTAAGTACGACCGGT 503  
DB 181 GCTACGCGCGGCGCGCTCTATTACGGGCTCATATGCGGCTGAGTAAGTACGACCGGT 240  
QY 504 GTTACCATCTACCTGGGACCTGGCGGCTGTGGTGGTCTGTTGGGCGGAGGCTGT 563  
DB 241 GTTACCATCTACCTGGGACCTGGCGGCTGTGGTGGTCTGTTGGGCGGAGGCTGT 300  
QY 564 GCGGAGGCGCTGGGAGGTGAGGCTGAGGAGTTACGCGCGCGGGAACCTGATGCT 623  
DB 301 GCGGAGGCGCTGGGAGGTGAGGCTGAGGAGTTACGCGCGCGGGAACCTGATGCT 360  
QY 624 GGAAGGACTTTTGTATGCGCATGGGTTTCTTCTTCCAAACGGGAGCGGTGAGGAGCT 683  
DB 361 GGAAGGACTTTTGTATGCGCATGGGTTTCTTCTTCCAAACGGGAGCGGTGAGGAGCT 420  
QY 684 GAGGAAGTTTACCATGCTTGTCTGCGGACCTGGGATGGGAGAGGAGGCGAGGA 743  
DB 421 GAGGAAGTTTACCATGCTTGTCTGCGGACCTGGGATGGGAGAGGAGGCGAGGA 480  
QY 744 GCTGATCCAGGGGAGGCGCGGTGTCTGTTGGAGACATTCCAGGGGACAGAGAGCGCCC 803  
DB 481 GCTGATCCAGGGGAGGCGCGGTGTCTGTTGGAGACATTCCAGGGGACAGAGAGCGCCC 540  
QY 804 ATTGATGCTGCTGCTGCTGGCCCGAGGCACTCCAAAGCTAGTGTCTTCTCTCTT 863  
DB 541 ATTGATGCTGCTGCTGCTGGCCCGAGGCACTCCAAAGCTAGTGTCTTCTCTCTT 600  
QY 864 TGSCCTCCGCTTCTCTATGAGGATAAGGAGTTCCAGGCGGTGTCGGGCGAGCTGGTG 923  
DB 601 TGSCCTCCGCTTCTCTATGAGGATAAGGAGTTCCAGGCGGTGTCGGGCGAGCTGGTG 660  
QY 924 TACCTGTGCGGAGTCACTCCAGGGGGTCCAGACTACGAGATTTCTCTGGTTCCT 983  
DB 661 TACCTGTGCGGAGTCACTCCAGGGGGTCCAGACTACGAGATTTCTCTGGTTCCT 720  
QY 984 GCGGCGGCTGCCAGGCGGCGGCAAGGAGCTCTTCCACCGCTCAGCACCTTGGCTGCTT 1043  
DB 721 GCGGCGGCTGCCAGGCGGCGGCAAGGAGCTCTTCCACCGCTCAGCACCTTGGCTGCTT 780



QY 1044 CACAGTCCGGCAGGTGCAGAGCAACAGGGGAACTGTGATGCTTCGGGCCCCGACAGTGA 1103  
Db 781 CACAGTCCGGCAGGTGCAGAGCAACAGGGGAACTGTGATGCTTCGGGCCCCGACAGTGA 840  
QY 1104 CTTTGTGATGCTTCCTGCTGAAGATGGCAGAGGAAACAAACCCAGGCAACAGAAAT 1163  
Db 841 CTTTGTGATGCTTCCTGCTGAAGATGGCAGAGGAAACAAACCCAGGCAACAGAAAT 900  
QY 1164 CACCAACAGAACATGCTGATGACAGTCAATTAATTTGCTGTTGCTGGGACGATGACGGT 1223  
Db 901 CACCAACAGAACATGCTGATGACAGTCAATTAATTTGCTGTTGCTGGGACGATGACGGT 960  
QY 1224 CAGCACACAGTCCGGCTATACCTCCCTGCTCCTGATGAAATACCTCATGTCCAAAGTG 1283  
Db 961 CAGCACACAGTCCGGCTATACCTCCCTGCTCCTGATGAAATACCTCATGTCCAAAGTG 1020  
QY 1284 GGTACGTGAGGAGCTGAATCCGGGAGCTGGGGGCTGGCCAGGCACCAAGCCTAGGGGACCG 1343  
Db 1021 GGTACGTGAGGAGCTGAATCCGGGAGCTGGGGGCTGGCCAGGCACCAAGCCTAGGGGACCG 1080  
QY 1344 TACCCGCTTCCTTACACCGACGGGTTCTGCATGAGCGCAGCGGCTGTGGCGCTGGT 1403  
Db 1081 TACCCGCTTCCTTACACCGACGGGTTCTGCATGAGCGCAGCGGCTGTGGCGCTGGT 1140  
QY 1404 GCCCATGGGAATACCCCGCACCTCATGCGGACCAACCGCTTCGAGGGGTACACCTGCC 1463  
Db 1141 GCCCATGGGAATACCCCGCACCTCATGCGGACCAACCGCTTCGAGGGGTACACCTGCC 1200  
QY 1464 CCAGGGCAACGAGGCTTTCCTCCCTCCTTGGCTCCATCTGCATGACCCCAACATCTTCAA 1523  
Db 1201 CCAGGGCAACGAGGCTTTCCTCCCTCCTTGGCTCCATCTGCATGACCCCAACATCTTCAA 1260  
QY 1524 GCACCCAGAGAGTTTAAACCGACCGTTTCCCTGGATGCAGATGGAACGGTTTCAGGAAGCA 1583  
Db 1261 GCACCCAGAGAGTTTAAACCGACCGTTTCCCTGGATGCAGATGGAACGGTTTCAGGAAGCA 1320  
QY 1584 TGAGGCGTTCCTGCCTTCCTTAGGAGAGCGTGTCTGCTTGGAGAGGCGCTGGCAAA 1643  
Db 1321 TGAGGCGTTCCTGCCTTCCTTAGGAGAGCGTGTCTGCTTGGAGAGGCGCTGGCAAA 1380  
QY 1644 AGCGAGCTTCCTCCTTTCACCAACATCCTCAAGACCTTCTCCCTGGAGAGCGCGTG 1703  
Db 1381 AGCGAGCTTCCTCCTTTCACCAACATCCTCAAGACCTTCTCCCTGGAGAGCGCGTG 1440  
QY 1704 CCGCCCGGACACCTGAGCTCAAGCCACCGTCAAGTGGCGCTTTTCAACATTCCTCCAGC 1763  
Db 1441 CCGCCCGGACACCTGAGCTCAAGCCACCGTCAAGTGGCGCTTTTCAACATTCCTCCAGC 1500  
QY 1764 CTTCCAGCTGCAAGTCCGTCCTCCTGACCTTCACTCCACACGACAGACAGATGAAGGA 1823  
Db 1501 CTTCCAGCTGCAAGTCCGTCCTCCTGACCTTCACTCCACACGACAGACAGATGAAGGA 1560  
QY 1824 GGCAACTTGAAGTGGTGGTGCCAGGCGTGCCTCCAGCCTCAACAGTGGGCATGGA 1883  
Db 1561 GGCAACTTGAAGTGGTGGTGCCAGGCGTGCCTCCAGCCTCAACAGTGGGCATGGA 1620  
QY 1884 CAGGGTTAATGCTCTCAGAGTGTACATGAGGAGGACCAATTACAGCCTGCAAGTTGT 1943  
Db 1621 CAGGGTTAATGCTCTCAGAGTGTACATGAGGAGGACCAATTACAGCCTGCAAGTTGT 1680  
QY 1944 TTTCCGAGTCTGTCCACGCGCCACACGCTCACTTGACTCATGCTCATGATGACAA 2003  
Db 1681 TTTCCGAGTCTGTCCACGCGCCACACGCTCACTTGACTCATGCTCATGATGACAA 1740  
QY 2004 CCGCACACCCATACAACTACAAGGGCCCAAGCAACTGCTGGGTAGCTTTCCACAG 2063  
Db 1741 CCGCACACCCATACAACTACAAGGGCCCAAGCAACTGCTGGGTAGCTTTCCACAG 1800  
QY 2064 ACATAATATAGTCCATCTGAAATCAACAGCATAGCCAGGTAACCCAACTCCCTC 2123  
Db 1801 ACATAATATAGTCCATCTGAAATCAACAGCATAGCCAGGTAACCCAACTCCCTC 1860

QY 2124 GGATCTGAGGCCACACAGTGGGAGTCTGGCTGTGCTTCAAGGCCACAGAAACGGCCA 2183  
Db 1861 GGATCTGAGGCCACACAGTGGGAGTCTGGCTGTGCTTCAAGGCCACAGAAACGGCCA 1920  
QY 2184 CACATGTTACAGCTCAACAGCCCTCTCCATTTCATTCGAATCTTCAGTGTCCCTGCTCCCT 2243  
Db 1921 CACATGTTACAGCTCAACAGCCCTCTCCATTTCATTCGAATCTTCAGTGTCCCTGCTCCCT 1980  
QY 2244 GGTGCTTGGCACAGGGAACAGCATGCCCCCTTCGCGGTTCATGCCACCCAGAGACTGTGCG 2303  
Db 1981 GGTGCTTGGCACAGGGAACAGCATGCCCCCTTCGCGGTTCATGCCACCCAGAGACTGTGCG 2040  
QY 2304 TGTATATGGCCCCCAACTCATGCTCCCTCTCTTGGCTACACACTCTCCAGCCTGTGACC 2363  
Db 2041 TGTATATGGCCCCCAACTCATGCTCCCTCTCTTGGCTACACACTCTCCAGCCTGTGACC 2100  
QY 2364 ACCGATGTCACACACACCCCAACACTTGTCCACACAGCTACCCAGTACGACATCGTCC 2423  
Db 2101 ACCGATGTCACACACACCCCAACACTTGTCCACACAGCTACCCAGTACGACATCGTCC 2160  
QY 2424 TGGTCTCCCGAGAGTATCTTCCCACTGAGACAACCGCCCCCACAGAGGACAGTCCCCAG 2483  
Db 2161 TGGTCTCCCGAGAGTATCTTCCCACTGAGACAACCGCCCCCACAGAGGACAGTCCCCAG 2220  
QY 2484 CCACCTCTGCAACTGAGCGCTCAGTCAACCCCTTTTAAAGCACCTGATTTACAAATG 2543  
Db 2221 CCACCTCTGCAACTGAGCGCTCAGTCAACCCCTTTTAAAGCACCTGATTTACAAATG 2280  
QY 2544 CAAACACATCTGGGTCTGCGATTATGACACAGAGACTTTGGACATACGAGGACCTTCAGAC 2603  
Db 2281 CAAACACATCTGGGTCTGCGATTATGACACAGAGACTTTGGACATACGAGGACCTTCAGAC 2340  
QY 2604 CGAGGAAACCTGCCCCAACCCCAACACGCTTATGTAACACGCTGGAAGCGGCCCT 2663  
Db 2341 CGAGGAAACCTGCCCCAACCCCAACACGCTTATGTAACACGCTGGAAGCGGCCCT 2400  
QY 2664 GCTGCCCTCCACACACATACACACTCACTGATCTACAGCCCTGTTGGCGTTCAGAG 2723  
Db 2401 GCTGCCCTCCACACACATACACACTCACTGATCTACAGCCCTGTTGGCGTTCAGAG 2460  
QY 2724 TCCCACTAGACCCAGTGGAGGGTTAGAGACCAAGTACGAGGCGAGTTCCCAATTCACC 2783  
Db 2461 TCCCACTAGACCCAGTGGAGGGTTAGAGACCAAGTACGAGGCGAGTTTCCAATTCACC 2520  
QY 2784 CTGTGAGGAGTACGCGGATCTGACGTTCTTGTGATTAAGGTCGCGCTTCGGGAAT 2843  
Db 2521 CTGTGAGGAGTACGCGGATCTGACGTTCTTGTGATTAAGGTCGCGCTTCGGGAAT 2580  
QY 2844 AAAGTTTGTCTGCGCTTTAGCCTAAAAAANAAAAA 2886  
Db 2581 AAAGTTTGTCTGCGCTTTAGCCTAAAAAANAAAAA 2623

RESULT 9  
ADA45582

ID ADA45582 standard; cDNA; 2623 BP.

XX ADA45582;

XX 20-NOV-2003 (first entry)

XX Novel human secreted and transmembrane protein PRO1906 cDNA.

XX Human; secreted and transmembrane protein; PRO; gene; ss;  
XX Tumour necrosis factor alpha release; TNF-alpha release;  
XX Glucose uptake modulator; FFA uptake modulator;  
XX Cell proliferation stimulator; cell differentiation stimulator;  
XX cell differentiation inhibitor; cytokine release stimulator; tumour;  
XX lung tumour; colon tumour; breast tumour; prostate tumour; rectal tumour;  
XX cervical tumour; liver tumour; chromosome mapping; gene mapping;  
XX gene therapy; chromosome identification; chromosome marker.  
OS Homo sapiens.

XX	US2003022328-A1.	PR	08-NOV-2000;	200WO-US030952.
PN		PR	10-NOV-2000;	200WO-US030873.
XX	30-JAN-2003.	PR	01-DEC-2000;	200WO-US032678.
PD		PR	20-DEC-2000;	200WO-US034956.
XX		PR	28-FEB-2001;	2001US-00796498.
XX		PR	28-FEB-2001;	2001WO-US006520.
XX		PR	01-MAR-2001;	2001WO-US006666.
PR	31-MAR-1997;	PR	09-MAR-2001;	2001US-00802706.
PR	12-JUN-1998;	PR	14-MAR-2001;	2001US-00806689.
PR	14-JUL-1998;	PR	22-MAR-2001;	2001US-00816744.
PR	28-AUG-1998;	PR	05-APR-2001;	2001US-00828366.
PR	10-SEP-1998;	PR	10-MAY-2001;	2001US-00854208.
PR	14-SEP-1998;	PR	10-MAY-2001;	2001US-00854280.
PR	14-SEP-1998;	PR	18-MAY-2001;	2001US-00860216.
PR	14-SEP-1998;	PR	25-MAY-2001;	2001US-00866028.
PR	16-SEP-1998;	PR	25-MAY-2001;	2001US-00866034.
PR	17-SEP-1998;	PR	25-MAY-2001;	2001WO-US017092.
PR	07-OCT-1998;	PR	01-JUN-2001;	2001US-00872035.
PR	29-OCT-1998;	PR	01-JUN-2001;	2001WO-US017800.
PR	20-NOV-1998;	PR	05-JUN-2001;	2001US-00874503.
PR	01-DEC-1998;	PR	14-JUN-2001;	2001US-00882636.
PR	05-JAN-1999;	PR	19-JUN-2001;	2001US-00886342.
PR	08-MAR-1999;	PR	20-JUN-2001;	2001WO-US019692.
PR	10-MAR-1999;	PR	21-JUN-2001;	2001US-00887879.
PR	20-APR-1999;	PR	22-JUN-2001;	2001WO-US020116.
PR	14-MAY-1999;	PR	29-JUN-2001;	2001WO-US021066.
PR	02-JUN-1999;	PR	09-JUL-2001;	2001WO-US021735.
PR	01-SEP-1999;	PR	18-JUL-2001;	2001US-00908827.
PR	08-SEP-1999;	PR	06-AUG-2001;	2001US-00924419.
PR	13-SEP-1999;	PR	09-AUG-2001;	2001US-00927796.
PR	15-SEP-1999;	PR	16-AUG-2001;	2001US-00931836.
PR	05-OCT-1999;	PR	19-DEC-2001;	2001US-00028072.
PR	29-NOV-1999;	XX		
PR	30-NOV-1999;	PA	(GETH ) GENENTECH INC.	
PR	01-DEC-1999;	XX		
PR	01-DEC-1999;	PI	Baker KP, Beresini M, Deforge L, Desnoyers L, Filvaroff E, Gao W;	
PR	01-DEC-1999;	PI	Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;	
PR	02-DEC-1999;	PI	Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;	
PR	02-DEC-1999;	XX		
PR	02-DEC-1999;	DR	WPI; 2003-584997/55.	
PR	16-DEC-1999;	DR	P-PSDB; ADA45583.	
PR	16-DEC-1999;	XX		
PR	20-DEC-1999;	PT	Novel secreted and transmembrane polypeptide for modulating biological	
PR	20-DEC-1999;	PT	activity of cell expressing the polypeptide, identifying agonists or	
PR	22-DEC-1999;	PT	antagonists of polypeptide, and as molecular weight markers.	
PR	30-DEC-1999;	XX		
PR	05-JAN-2000;	XX	Claim 2; Fig 63; 659pp; English.	
PR	05-JAN-2000;	XX		
PR	06-JAN-2000;	CC	The invention describes 305 nucleic acids encoding PRO (secreted and	
PR	06-JAN-2000;	CC	transmembrane) polypeptides (I). (I) is useful for stimulating the	
PR	11-FEB-2000;	CC	release of TNF-alpha from human blood, for modulating the uptake of	
PR	18-FEB-2000;	CC	glucose or FFA by skeletal muscle cells or adipocyte cells, for	
PR	18-FEB-2000;	CC	stimulating the proliferation or differentiation of chondrocyte cells,	
PR	22-FEB-2000;	CC	for stimulating the proliferation of or gene expression in pericyte	
PR	24-FEB-2000;	CC	cells, for stimulating the release of proteoglycans from cartilage, for	
PR	24-FEB-2000;	CC	stimulating the proliferation of inner ear utricular supporting cells,	
PR	01-MAR-2000;	CC	for stimulating the proliferation of T-lymphocyte cells, for stimulating	
PR	02-MAR-2000;	CC	the release of a cytokine from PBMC cells, for inhibiting the binding of	
PR	02-MAR-2000;	CC	A-peptide to factor VIIa, for inhibiting the differentiation of adipocyte	
PR	10-MAR-2000;	CC	cells, for stimulating proliferation of endothelial cells, for detecting	
PR	15-MAR-2000;	CC	the presence of tumour in a mammal. The tumour is lung, colon, breast,	
PR	20-MAR-2000;	CC	prostate, rectal, cervical or liver tumour. The oligonucleotide probes	
PR	21-MAR-2000;	CC	are useful for isolating genomic and cDNA nucleotide sequences or	
PR	21-MAR-2000;	CC	antisense probes. (I) is also useful as therapeutic agent. PRO is useful	
PR	17-MAY-2000;	CC	in assays to identify other proteins or molecules involved in binding	
PR	22-MAY-2000;	CC	interaction. A polynucleotide (ii) encoding (I) is useful in chromosome	
PR	30-MAY-2000;	CC	and gene mapping, in generation of antisense RNA and DNA, in the	
PR	02-JUN-2000;	CC	preparation of PRO polypeptide, for generating transgenic animals or	
PR	28-JUL-2000;	CC	knockout animals which in turn are useful in the development and	
PR	11-AUG-2000;	CC	screening of therapeutically useful reagents, in gene therapy, for	
PR	23-AUG-2000;	CC	chromosome identification, as chromosome marker, and for generating	
PR	24-AUG-2000;	CC	probes. An anti-(I)-antibody is useful in diagnostic assays for PRO, e.g.	



QY	2304	TGTCATGGCCCCCAACTCATGCTCCCTCTCTTGGCTACACCACTCTCCAGGCTGTGACC	2363	PR	14-JUL-1998;	98WO-US014552.
				PR	28-AUG-1998;	98WO-US017898.
				PR	10-SEP-1998;	98WO-US018824.
Db	2041	TGTCATGGCCCCAACTCATGCTCCCTCTCTTGGCTACACCACTCTCCAGGCTGTGACC	2100	PR	14-SEP-1998;	98WO-US019093.
				PR	14-SEP-1998;	98WO-US019094.
QY	2364	ACCGATGTCCACACACACCCCAACCACTTGTCCACACAGCTACCCAGTACGACATCGTCC	2423	PR	14-SEP-1998;	98WO-US019177.
				PR	16-SEP-1998;	98WO-US019330.
Db	2101	ACCGATGTCCACACACACCCCAACCACTTGTCCACACAGTACCCAGTACACATCGTCC	2160	PR	17-SEP-1998;	98WO-US019437.
				PR	07-OCT-1998;	98WO-US021141.
QY	2424	TGGCTCCCGAGATATCTTCCCACTGAGACACGCGCGCCCCACAGGGACAGTCCCCAG	2483	PR	29-OCT-1998;	98WO-US022991.
				PR	29-OCT-1998;	98WO-US022992.
Db	2161	TGGCTCCCGAGATATCTTCCCACTGAGACACGCGCGCCCCACAGGGACAGTCCCCAG	2220	PR	20-NOV-1998;	98WO-US024855.
				PR	01-DEC-1998;	98WO-US025108.
QY	2484	CAACCTCTGCAACTGAGCCCTCAGTCACCCCTTTTAAAGCACCCCTGATTTACCAAAATG	2543	PR	05-JAN-1999;	99WO-US000106.
				PR	08-MAR-1999;	99WO-US005028.
Db	2221	CAACCTCTGCAACTGAGCCCTCAGTCACCCCTTTTAAAGCACCCCTGATTTACCAAAATG	2280	PR	10-MAR-1999;	99WO-US005190.
				PR	20-APR-1999;	99WO-US008615.
QY	2544	CAAAACATCTGGGTCTGGATTATGCAACAGAGACTTTTGGACATACGAGGACCCCTCAGAC	2603	PR	14-MAY-1999;	99WO-US010733.
				PR	02-JUN-1999;	99WO-US012252.
Db	2281	CAAAACATCTGGGTCTGGATTATGCAACAGAGACTTTTGGACATACGAGGACCCCTCAGAC	2340	PR	01-SEP-1999;	99WO-US020111.
				PR	08-SEP-1999;	99WO-US020594.
QY	2604	CGGAGAACACCTGCCCAACCCCAACACGTGCTTATGTAACCACTGTAAGCGGGCCCT	2663	PR	13-SEP-1999;	99WO-US020944.
				PR	15-SEP-1999;	99WO-US021090.
Db	2341	CGGAGAACACCTGCCCAACCCCAACACGTGCTTATGTAACCACTGTAAGCGGGCCCT	2400	PR	15-SEP-1999;	99WO-US021547.
				PR	05-OCT-1999;	99WO-US023089.
QY	2664	GCTGCCCTCCACACACATACACACTCACTGATCTACAGCCCTGTTTCGGCGTCAGAG	2723	PR	29-NOV-1999;	99WO-US028214.
				PR	30-NOV-1999;	99WO-US028313.
Db	2401	GCTGCCCTCCACACACATACACACTCACTGATCTACAGCCCTGTTTCGGCGTCAGAG	2460	PR	30-NOV-1999;	99WO-US028409.
				PR	01-DEC-1999;	99WO-US028301.
QY	2724	TCCCACTAGACCCAGTGAAGGGTTAGAGACCAAGTAGGGGCCAGTTTCCAATTCACC	2783	PR	01-DEC-1999;	99WO-US028634.
				PR	02-DEC-1999;	99WO-US028551.
Db	2461	TCCCACTAGACCCAGTGAAGGGTTAGAGACCAAGTAGGGGCCAGTTTCCAATTCACC	2520	PR	02-DEC-1999;	99WO-US028564.
				PR	12-DEC-1999;	99WO-US028565.
QY	2784	CTGTAGGGAGTAGCCGAGCTGAGTTCCTTGTGACTTAAGGTCGGCTTGGGAATT	2843	PR	16-DEC-1999;	99WO-US030095.
				PR	20-DEC-1999;	99WO-US030911.
Db	2521	CTGTAGGGAGTAGCCGAGCTGAGTTCCTTGTGACTTAAGGTCGGCTTGGGAATT	2580	PR	22-DEC-1999;	99WO-US030999.
				PR	22-DEC-1999;	99WO-US030720.
QY	2844	AAAGTTTGTCTTGGCTTTAGCTTAAAAAATAAAAAAAAAA 2886		PR	30-DEC-1999;	99WO-US031243.
				PR	05-JAN-2000;	2000WO-US000219.
Db	2581	AAAGTTTGTCTTGGCTTTAGCTTAAAAAATAAAAAAAAAA 2623		PR	06-JAN-2000;	2000WO-US000277.
				PR	11-FEB-2000;	2000WO-US003376.
				PR	18-FEB-2000;	2000WO-US003565.
				PR	18-FEB-2000;	2000WO-US004341.
				PR	22-FEB-2000;	2000WO-US004342.
				PR	24-FEB-2000;	2000WO-US004414.
				PR	24-FEB-2000;	2000WO-US004914.
				PR	01-MAR-2000;	2000WO-US005004.
				PR	02-MAR-2000;	2000WO-US005601.
				PR	02-MAR-2000;	2000WO-US005746.
				PR	10-MAR-2000;	2000WO-US005841.
				PR	15-MAR-2000;	2000WO-US006319.
				PR	20-MAR-2000;	2000WO-US006884.
				PR	21-MAR-2000;	2000WO-US007377.
				PR	30-MAR-2000;	2000WO-US007532.
				PR	17-MAY-2000;	2000WO-US008439.
				PR	22-MAY-2000;	2000WO-US013705.
				PR	30-MAY-2000;	2000WO-US014042.
				PR	02-JUN-2000;	2000WO-US014941.
				PR	28-JUL-2000;	2000WO-US015264.
				PR	11-AUG-2000;	2000WO-US020710.
				PR	23-AUG-2000;	2000WO-US022031.
				PR	24-AUG-2000;	2000WO-US023522.
				PR	08-NOV-2000;	2000WO-US023328.
				PR	10-NOV-2000;	2000WO-US030952.
				PR	01-DEC-2000;	2000WO-US030873.
				PR	20-DEC-2000;	2000WO-US032678.
				PR	20-DEC-2000;	2000US-00747259.
				PR	28-FEB-2001;	2000WO-US034956.
				PR	28-FEB-2001;	2001US-00796498.
				PR	01-MAR-2001;	2001WO-US006520.
				PR	09-MAR-2001;	2001WO-US006666.
				PR		2001US-00802706.

RESULT 10  
ADA76013  
ID ADA76013 standard; cDNA; 2623 BP.  
XX  
AC ADA76013;  
XX  
DT 20-NOV-2003 (first entry)  
XX  
DE Human PRO polynucleotide #32.  
XX  
KW Human; gene; ss; PRO; secreted polypeptide; transmembrane polypeptide;  
KW tumour necrosis factor-alpha; TNF-alpha; chondrocyte cell; tumour;  
KW cancer; adrenal; lung; colon; breast; prostate; rectum; kidney; cervix;  
KW liver; microvascular endothelial cell; glucose; FFA;  
KW skeletal muscle cell; adipocyte cell; pericyte cell;  
KW inner ear utricular supporting cell; T-lymphocyte cell;  
KW endothelial cell tube formation; bone disorder; cartilage disorder;  
KW sports injury; proteoglycan; articular cartilage defect; osteoarthritis;  
KW rheumatoid arthritis; haemoglobin-associated disorder thalassaemia;  
KW immune system cell infiltration.  
XX  
OS Homo sapiens.  
XX  
PN US2003073212-A1.  
XX  
PD 17-APR-2003.  
XX  
PF 16-APR-2002; 2002US-00123903.  
XX  
PR 31-MAR-1997; 97WO-US005230.  
PR 12-JUN-1998; 98WO-US012456.



Db 961 CAGCACACGGTGGGCTATACCTCTCTCTGATGAATAACCTCATGTCCAAAGTG 1020  
Qy 1284 GGTAGTGAAGAGCTGAATCGGAGCTGGGGCTGGCCAGGACCAAGCCTTAGGGAGCG 1343  
Db 1021 GGTACGTGAGGAGCTGAATCGGAGCTGGGGCTGGCCAGGACCAAGCCTTAGGGAGCG 1080  
Qy 1344 TACCGGCTCCCTTACCGAGCGCGGTCTGTGATGAGGCGGAGCGGCTGCTGGCGCTGGT 1403  
Db 1081 TACCGGCTCCCTTACCGAGCGCGGTCTGTGATGAGGCGGAGCGGCTGCTGGCGCTGGT 1140  
Qy 1404 GCCATGGGAAATACCCCGACCCCTCATGCGGACCAACCGCTTCCGAGGGTACACCTGCG 1463  
Db 1141 GCCATGGGAAATACCCCGACCCCTCATGCGGACCAACCGCTTCCGAGGGTACACCTGCG 1200  
Qy 1464 CCAGGCGAGGAGGTCTTCCCGCTCTTGGCTTCGATGCGATGACCCCAACATCTTCAA 1523  
Db 1201 CCAGGCGAGGAGGTCTTCCCGCTCTTGGCTTCGATGCGATGACCCCAACATCTTCAA 1260  
Qy 1524 GCACCCAGAAGAGTTCAACCCAGACCGTTTCTGTGATGCAGATGGACGGTTTCAGGAAGCA 1583  
Db 1261 GCACCCAGAAGAGTTCAACCCAGACCGTTTCTGTGATGCAGATGGACGGTTTCAGGAAGCA 1320  
Qy 1584 TGAGCGTTCTGGCGCTTCTCTTAGGGAAGCGTGTCTGCCCTTTCAGATGCGGCGCTGGCAAA 1643  
Db 1321 TGAGCGTTCTGGCGCTTCTCTTAGGGAAGCGTGTCTGCCCTTTCAGATGCGGCGCTGGCAAA 1380  
Qy 1644 AGCGAGCTCTTCTCTTCCACACCATCTTCAAGCGCTTCTCCCTGGAGAGCCGCTG 1703  
Db 1381 AGCGAGCTCTTCTCTTCCACACCATCTTCAAGCGCTTCTCCCTGGAGAGCCGCTG 1440  
Qy 1704 CCGCGCGGACACCTGAGCGCTCAAGCCCAACCGTTCAGTGGCGCTTTCACATTTCCCGCAGC 1763  
Db 1441 CCGCGCGGACACCTGAGCGCTCAAGCCCAACCGTTCAGTGGCGCTTTCACATTTCCCGCAGC 1500  
Qy 1764 CTTTCAGCTGGAAGTCCGTCCTCACTGACCTTCAAGCGCTTTCACATTTTCAGGAAGCA 1823  
Db 1501 CTTTCAGCTGGAAGTCCGTCCTCACTGACCTTCAAGCGCTTTCACATTTTCAGGAAGCA 1560  
Qy 1824 GGCACCTTGAAGTGGTGGTGGCCAGGAGCGTGGCTCCAGCTCAACAGTGGGATGGA 1883  
Db 1561 GGCACCTTGAAGTGGTGGTGGCCAGGAGCGTGGCTCCAGCTCAACAGTGGGATGGA 1620  
Qy 1884 CAGGGTTAATGTCTCCAGAGTGTACACTGCAGGCGACCAATTTACACGCTTCGAGTTGT 1943  
Db 1621 CAGGGTTAATGTCTCCAGAGTGTACACTGCAGGCGACCAATTTACACGCTTCGAGTTGT 1680  
Qy 1944 TTTCCGGAGTGTGTCCACGGGCCACACGCTCACTTGACTCATGTCTGAAGATGACAA 2003  
Db 1681 TTTCCGGAGTGTGTCCACGGGCCACACGCTCACTTGACTCATGTCTGAAGATGACAA 1740  
Qy 2004 CCGCACCCCATACACAACTCAAGGGGCCACAAAGCACTGTGGGTTAGCTTTCCACAG 2063  
Db 1741 CCGCACCCCATACACAACTCAAGGGGCCACAAAGCACTGTGGGTTAGCTTTCCACAG 1800  
Qy 2064 ACATAAATATAGTTCATCTGCAATCAACAGCATAGCCAGGTAAACCAACTCCCT 2123  
Db 1801 ACATAAATATAGTTCATCTGCAATCAACAGCATAGCCAGGTAAACCAACTCCCT 1860  
Qy 2124 GGATCTGCAGCCACACGCTGGGAGTGTGGCTGACCTTTCACAGCCACAGAAACGGCCA 2183  
Db 1861 GGATCTGCAGCCACACGCTGGGAGTGTGGCTGACCTTTCACAGCCACAGAAACGGCCA 1920  
Qy 2184 CACATGTTTACAGCTCACACGGCTCTCCATTCATCGAACTTCTCAGTGTCCCTGTCCT 2243  
Db 1921 CACATGTTTACAGCTCACACGGCTCTCCATTCATCGAACTTCTCAGTGTCCCTGTCCT 1980  
Qy 2244 GGTGCTGGCACAGGGAAACAGATGCCCTCCCGGGTCAAGCCACCAAGAGACTGTCCG 2303  
Db 1981 GGTGCTGGCACAGGGAAACAGATGCCCTCCCGGGTCAAGCCACCAAGAGACTGTCCG 2040  
Qy 2304 TGTCTATGGCCCCCACTCATGCTCTTGGCTTACACCACTCTCCAGCTGTGACC 2363  
Db 2041 TGTCTATGGCCCCCACTCATGCTCTTGGCTTACACCACTCTCCAGGCTGTGACC 2100

Qy 2364 ACCGATGTCCACACACCCCAACCACTTGTTCACACAGCTACCCAGCTACGACATCGTCC 2423  
Db 2101 ACCGATGTCCACACACCCCAACCACTTGTTCACACAGCTACCCAGCTACGACATCGTCC 2160  
Qy 2424 TGGTCTCCACAGATATCTTCCCACTGAGACAGCGCGGCCCCACAGAGGCACAGTCCCCAG 2483  
Db 2161 TGGTCTCCACAGATATCTTCCCACTGAGACAGCGCGGCCCCACAGAGGCACAGTCCCCAG 2220  
Qy 2484 CCACCTCTGCAACTGACGCGCTCAGTCAACCCCTTTTAAAGCAACCCCTGATCTACCAATG 2543  
Db 2221 CCACCTCTGCAACTGACGCGCTCAGTCAACCCCTTTTAAAGCAACCCCTGATCTACCAATG 2280  
Qy 2544 CAAACACATCTGGGTCTCGATTATGCACAGAGCTTTGGACATACGAGGACCCCTCAGAC 2603  
Db 2281 CAAACACATCTGGGTCTCGATTATGCACAGAGCTTTGGACATACGAGGACCCCTCAGAC 2340  
Qy 2604 CGAGGAAACACTGCGCCCAACCCCAACAGCTGCTTATGTAACCAACGCGGAAAGCGGCCCT 2663  
Db 2341 CGAGGAAACACTGCGCCCAACCCCAACAGCTGCTTATGTAACCAACGCGGAAAGCGGCCCT 2400  
Qy 2664 GCTGCGCTTCCACACACATACACACTCAGTATCTACAGCCCTGTTTCGGCGTCAGAG 2723  
Db 2401 GCTGCGCTTCCACACACATACACACTCAGTATCTACAGCCCTGTTTCGGCGTCAGAG 2460  
Qy 2724 TCCCACTAGACCCCACTGGAAGGGGTTAGAGACCAAGTAGGGGCGCAGTTTCCAAATTCACC 2783  
Db 2461 TCCCACTAGACCCCACTGGAAGGGGTTAGAGACCAAGTAGGGGCGCAGTTTCCAAATTCACC 2520  
Qy 2784 CTGTGAGGAGTGAAGCGGATCTGAGCTTCTGTGACTTAAAGGTCGCGCTTGGGAAT 2843  
Db 2521 CTGTGAGGAGTGAAGCGGATCTGAGCTTCTGTGACTTAAAGGTCGCGCTTGGGAAT 2580  
Qy 2844 AAGTTTGTCTGCGCTTTAGCCTTAAAGGTCGCGCTTAAAGGTCGCGCTTGGGAAT 2886  
Db 2581 AAGTTTGTCTGCGCTTTAGCCTTAAAGGTCGCGCTTAAAGGTCGCGCTTGGGAAT 2623

## RESULT 11

ADAL8663  
ID ADAL8663 standard; cDNA; 2623 BP.  
XX AC ADAL8663;  
XX DT 20-NOV-2003 (first entry)  
XX DE Human PRO polynucleotide #32.  
XX KW Human; gene; ss; PRO; secreted polypeptide; transmembrane polypeptide; tumour necrosis factor-alpha; TNF-alpha; blood; chondrocyte cell; lung; colon; breast; prostate; rectum; cervix; liver; tumour; cancer;  
KW glucose uptake; PFA; adipocyte cell; pericyte cell; proteoglycan;  
KW cartilage; inner ear utricular supporting cell; cytokine; A-peptide;  
KW factor VIIA; endothelial cell.  
XX OS Homo sapiens.  
XX PN US2003054517-A1.  
XX PD 20-MAR-2003.  
XX PF 08-MAY-2002; 2002US-00141755.  
XX PR 31-MAR-1997; 97WO-US005230.  
PR 12-JUN-1998; 98WO-US012456.  
PR 14-JUL-1998; 98WO-US014552.  
PR 28-AUG-1998; 98WO-US017888.  
PR 10-SEP-1998; 98WO-US018824.  
PR 14-SEP-1998; 98WO-US019093.  
PR 14-SEP-1998; 98WO-US019094.  
PR 14-SEP-1998; 98WO-US019177.  
PR 16-SEP-1998; 98WO-US019330.  
PR 17-SEP-1998; 98WO-US019437.





QY 384 GSCCGAGGCCACCTGCCCCCGGGCCACAGCGCTACACCTGCTGGGAAACCTCTGCA 443  
DB 121 GSCCGAGGCCACCTGCCCCCGGGCCACAGCGCTACACCTGCTGGGAAACCTCTGCA 180  
QY 444 GCTACGGGCCGGGGCGCTGTATTCAGGGCTCATGCGGCTGAGTAAGAAGTACGGACCGGT 503  
DB 181 GCTACGGGCCGGGGCGCTGTATTCAGGGCTCATGCGGCTGAGTAAGAAGTACGGACCGGT 240  
QY 504 GTTACCATCTACCTGGAGACCTGCGGCTGTGTGTGCTGCTGCTGGTGGCAGAGGCTGT 563  
DB 241 GTTACCATCTACCTGGAGACCTGCGGCTGTGTGTGCTGCTGCTGGTGGCAGAGGCTGT 300  
QY 564 GCGGAGGCGCTGGAGGTCAGGCTGAGAGTTCAAGCGCTGAGCGGGAACCTGATGCT 623  
DB 301 GCGGAGGCGCTGGAGGTCAGGCTGAGAGTTCAAGCGCTGAGCGGGAACCTGATGCT 360  
QY 624 GGAAGGACTTTTGTATGGCCATGGGGTTTTCTTCTCAAACGGGAGCGGTGAGGAGCT 683  
DB 361 GGAAGGACTTTTGTATGGCCATGGGGTTTTCTTCTCAAACGGGAGCGGTGAGGAGCT 420  
QY 684 GAGGAAGTTTACCATGCTGTCTGCGGACCTGCGCATGGGAGCGGAGAGGCGAGGA 743  
DB 421 GAGGAAGTTTACCATGCTGTCTGCGGACCTGCGCATGGGAGCGGAGAGGCGAGGA 480  
QY 744 GCTGATCCAGGCGGAGGCGCGGTGTCTGTGTGAGACATTCAGGGGACAGAAAGCGCC 803  
DB 481 GCTGATCCAGGCGGAGGCGCGGTGTCTGTGTGAGACATTCAGGGGACAGAAAGCGCC 540  
QY 804 ATTGATCCCTCCTGCTGCTGCGCCAGGCCAACCCTCAAACGTAGTCTGTCTCCTCTT 863  
DB 541 ATTGATCCCTCCTGCTGCTGCGCCAGGCCAACCCTCAAACGTAGTCTGTCTCCTCTT 600  
QY 864 TGGCCTCGCTTCTCTATGAGTAAGAGTTCCAGGCGGTGCTCGGCGAGCTGGTG 923  
DB 601 TGGCCTCGCTTCTCTATGAGTAAGAGTTCCAGGCGGTGCTCGGCGAGCTGGTG 660  
QY 924 TACCTGCTGGAGTCAGCTCCAGGGGGTACAGCTACGAGATGTTCTCTGGTTCT 983  
DB 661 TACCTGCTGGAGTCAGCTCCAGGGGGTACAGCTACGAGATGTTCTCTGGTTCT 720  
QY 984 GCGGCCCTTCCAGGCCCCCAAGCAGTCTCTTCCACAGCTCAGCACTTGGCTGCCCTT 1043  
DB 721 GCGGCCCTTCCAGGCCCCCAAGCAGTCTCTTCCACAGCTCAGCACTTGGCTGCCCTT 780  
QY 1044 CACAGTCGGCAGGTGCAGCAGCACAGGGACCTGGATGCTTCGGGCCCGCAGTGGA 1103  
DB 781 CACAGTCGGCAGGTGCAGCAGCACAGGGACCTGGATGCTTCGGGCCCGCAGTGGA 840  
QY 1104 CTTTGTGATGCTTCTGCTGAAGATGGCAGGAGGAAACAAACCCAGGCACAGAAAT 1163  
DB 841 CTTTGTGATGCTTCTGCTGAAGATGGCAGGAGGAAACAAACCCAGGCACAGAAAT 900  
QY 1164 CACCAACAGAACATGCTGATGACGTCAATTTATTTGCTGTTGCTGGGACGATGACGT 1223  
DB 901 CACCAACAGAACATGCTGATGACGTCAATTTATTTGCTGTTGCTGGGACGATGACGT 960  
QY 1224 CAGCACCGCTGGCTTACCTCTCTCTGATGAAATACCTCATGTCCAAAGTG 1283  
DB 961 CAGCACCGCTGGCTTACCTCTCTCTGATGAAATACCTCATGTCCAAAGTG 1020  
QY 1284 GGTACGTGAGGAGCTGAATCGGAGCTGGGGGTGGCCAGGCACCAAGCTTAGGGGACCG 1343  
DB 1021 GGTACGTGAGGAGCTGAATCGGAGCTGGGGGTGGCCAGGCACCAAGCTTAGGGGACCG 1080  
QY 1344 TACCCGCTTCCCTTACACGACGCGGTTCTGATGAGGCGCAGCGGCTGCTGCGCTGGT 1403  
DB 1081 TACCCGCTTCCCTTACACGACGCGGTTCTGATGAGGCGCAGCGGCTGCTGCGCTGGT 1140  
QY 1404 GCCCATGGGAATACCCCGACCTCATGCGACCAACCCGCTCCGAGGGTACACCTGCC 1463  
DB 1141 GCCCATGGGAATACCCCGACCTCATGCGACCAACCCGCTCCGAGGGTACACCTGCC 1200  
QY 1464 CCAGGGCAGGAGGTCTTCCCTCTCTGCTCCATCTCTGATGACCCCAACATCTTCAA 1523

DB 1201 CCAGGGCAGGAGGTCTTCCCTCTCTGGCTCCATCTCTGATGACCCCAACATCTTCAA 1260  
QY 1524 GCACCCAGAGAGTTCAACCCAGACGGTTTCTGTGATGAGATGACGGTTTCAAGGAGCA 1583  
DB 1261 GCACCCAGAGAGTTCAACCCAGACGGTTTCTGTGATGAGATGACGGTTTCAAGGAGCA 1320  
QY 1584 TGAGGCTTCTGCTTCTCTCTAGGGAAGGCTGTCTGCTTGGAGAGGCTTGGCAAA 1643  
DB 1321 TGAGGCTTCTGCTTCTCTCTAGGGAAGGCTGTCTGCTTGGAGAGGCTTGGCAAA 1380  
QY 1644 AGCGAGCTTCTTCTTCTTCCACCACTCTCAAGCCTTCTCTCTGGAGAGCCGCTG 1703  
DB 1381 AGCGAGCTTCTTCTTCTTCCACCACTCTCAAGCCTTCTCTCTGGAGAGCCGCTG 1440  
QY 1704 CCGCGGACACCTCTGAGCTCAAGCCACCGTCTGAGTGGCTTTCAACATTTCCCCAGC 1763  
DB 1441 CCGCGGACACCTCTGAGCTCAAGCCACCGTCTGAGTGGCTTTTCAACATTTCCCCAGC 1500  
QY 1764 CTTTCAGCTGCAAGTCCGTTCCACTGACCTTCACTCCACAGCGACAGATGAAGGAA 1823  
DB 1501 CTTTCAGCTGCAAGTCCGTTCCACTGACCTTCACTCCACAGCGACAGATGAAGGAA 1560  
QY 1824 GGCAACTTGGAAAGTGGTGGTCCCGAGGACGGTGTCTCAGCTCAACAGTGGGCTATGA 1883  
DB 1561 GGCAACTTGGAAAGTGGTGGTCCCGAGGACGGTGTCTCAGCTCAACAGTGGGCTATGA 1620  
QY 1884 CAGGGTTAATGTCTCCAGAGTGTACCTGACGAGCGCACATTTACCGCTTGCAGTGT 1943  
DB 1621 CAGGGTTAATGTCTCCAGAGTGTACCTGACGAGCGCACATTTACCGCTTGCAGTGT 1680  
QY 1944 TTTTCGGAGTCTGTCCACGGGCCACACGCTCACTTGACTCATGTCTGAAGATGCACAA 2003  
DB 1681 TTTTCGGAGTCTGTCCACGGGCCACACGCTCACTTGACTCATGTCTGAAGATGCACAA 1740  
QY 2004 CCGCACACCCATACAACTAACAGGCCCAAAAGCAACTGTCTGGTTAGCTTTCACAG 2063  
DB 1741 CCGCACACCCATACAACTAACAGGGCCACAAAGCAACTGTCTGGTTAGCTTTCACAG 1800  
QY 2064 ACATAAATATAGTCCATCTGCAATCAAGACATAGCCAGGTAAACCCCACTCCCT 2123  
DB 1801 ACATAAATATAGTCCATCTGCAATCAAGACATAGCCAGGTAAACCCCACTCCCT 1860  
QY 2124 GGATCTGACGCCACACGCTGGGAGTCTGCTGTGCTACCTTCACAAGCCACAGAAACGGCCA 2183  
DB 1861 GGATCTGACGCCACACGCTGGGAGTCTGCTGTGCTACCTTCACAAGCCACAGAAACGGCCA 1920  
QY 2184 CACATGTTACAGCTCAGCGCTCTCCATTCATCGAACTTCTCAGTCTCCTGTCCCT 2243  
DB 1921 CACATGTTACAGCTCAGCGCTCTCCATTCATCGAACTTCTCAGTCTCCTGTCCCT 1980  
QY 2244 GGTGCTGGCACAGGGAACAGATGCCCTCCGGGTCTATGCCACCCAGAGACTGTCCG 2303  
DB 1981 GGTGCTGGCACAGGGAACAGATGCCCTCCGGGTCTATGCCACCCAGAGACTGTCCG 2040  
QY 2304 TGTCTATGGCCCCAACTCATGTCTCTCTTTGGCTACCACTCTCCAGCTGTGAC 2363  
DB 2041 TGTCTATGGCCCCAACTCATGTCTCTCTTTGGCTACCACTCTCCAGCTGTGAC 2100  
QY 2364 ACCGATGTCCACACACCCCAACCACTTGTCCACAGCTACCCAGTACGATCGTCC 2423  
DB 2101 ACCGATGTCCACACACCCCAACCACTTGTCCACAGCTACCCAGTACGATCGTCC 2160  
QY 2424 TGGCTCCCGAGAGTATCTTCCCACTGAGACAGCCGCCCCCAACAGAGGACAGTCCCCAG 2483  
DB 2161 TGGCTCCCGAGAGTATCTTCCCACTGAGACAGCCGCCCCCAACAGAGGACAGTCCCCAG 2220  
QY 2484 CCACTCTGCAACTGAGCGCTCAGTCAACCCCTTTTAAAGACCCCTGATTTTCAACAAATG 2543  
DB 2221 CCACTCTGCAACTGAGCGCTCAGTCAACCCCTTTTAAAGACCCCTGATTTTCAACAAATG 2280  
QY 2544 CAAGAACATCTGGGTCTGGATTTATGACAGAGACTTTGGACATAGGAGACCTCAGAC 2603

Db 2281 CAACACATCTGGGTCTGGATTATGACACAGACCTTTGGACATAGAGGACCCCTCAGAC 2340  
Qy 2604 CGGAGGACACCTGCCAACCACCCACACGTCCTTATGTAACACAGTGGAAAGCGGCCCT 2663  
Db 2341 CGGAGGAACACCTGCCAACCACCCACACGTCCTTATGTAACACAGTGGAAAGCGGCCCT 2400  
Qy 2664 GCTGCCCTCCACACACACATACACACTCACTGATCTACAGCCCTGTTGCGGCTCAGAG 2723  
Db 2401 GCTGCCCTCCACACACACATACACACTCACTGATCTACAGCCCTGTTGCGGCTCAGAG 2460  
Qy 2724 TCCCACTAGACCAGTGGAGGGTTAGACACCAAGTAGGGGCCAGTTTCCAATTCAAC 2783  
Db 2461 TCCCACTAGACCAGTGGAGGGTTAGACACCAAGTAGGGGCCAGTTTCCAATTCAAC 2520  
Qy 2784 CTGTCAGGAGTGAAGCGGATCTACGTCCTTCTGATTAAGGTCGGCTTGGGAATT 2843  
Db 2521 CTGTCAGGAGTGAAGCGGATCTACGTCCTTCTGATTAAGGTCGGCTTGGGAATT 2580  
Qy 2844 AAAGTTTGTCTTCTGGCCTTTAGCCTAAAAAAGGAAAAA 2886  
Db 2581 AAAGTTTGTCTTCTGGCCTTTAGCCTAAAAAAGGAAAAA 2623

RESULT 12  
ADA61286  
ID ADA61286 standard; cDNA; 2623 BP.  
XX AC ADA61286;  
XX  
DT 20-NOV-2003 (first entry)  
XX  
DE Homo sapiens.  
XX  
KW Human; secreted and transmembrane protein; PRO; gene; ss;  
KW tumour necrosis factor alpha release; TNF-alpha release;  
KW glucose uptake modulator; FFA uptake modulator;  
KW cell proliferation stimulator; cell differentiation stimulator;  
KW cell differentiation inhibitor; cytokine release stimulator; tumour;  
KW lung tumour; colon tumour; breast tumour; prostate tumour; rectal tumour;  
KW cervical tumour; liver tumour; chromosome mapping; gene mapping;  
KW gene therapy; chromosome identification; chromosome marker.  
XX  
OS Novel.  
OS human.  
OS secreted.  
OS and.  
OS transmembrane.  
OS protein.  
OS PRO1906.  
OS CDNA.  
XX  
XX US2003049816-A1.  
XX  
PN 13-MAR-2003.  
PD  
PF 15-APR-2002; 2002US-00123262.  
XX  
XX 31-MAR-1997; 97WO-US005230.  
PR 12-JUN-1998; 98WO-US012456.  
PR 14-JUL-1998; 98WO-US014552.  
PR 28-AUG-1998; 98WO-US017888.  
PR 10-SEP-1998; 98WO-US018824.  
PR 14-SEP-1998; 98WO-US019093.  
PR 14-SEP-1998; 98WO-US019094.  
PR 14-SEP-1998; 98WO-US019177.  
PR 16-SEP-1998; 98WO-US019330.  
PR 17-SEP-1998; 98WO-US019437.  
PR 07-OCT-1998; 98WO-US021141.  
PR 29-OCT-1998; 98WO-US022991.  
PR 29-OCT-1998; 98WO-US022992.  
PR 20-NOV-1998; 98WO-US024855.  
PR 01-DEC-1998; 98WO-US025108.  
PR 05-JAN-1999; 99WO-US000106.  
PR 08-MAR-1999; 99WO-US005028.  
PR 10-MAR-1999; 99WO-US005190.  
PR 20-APR-1999; 99WO-US008615.  
PR 14-MAY-1999; 99WO-US010733.  
PR 01-JUN-1999; 99WO-US012252.  
PR 01-SEP-1999; 99WO-US020111.  
PR 08-SEP-1999; 99WO-US020594.  
PR 13-SEP-1999; 99WO-US020944.  
PR 15-SEP-1999; 99WO-US021090.  
PR 15-SEP-1999; 99WO-US021547.  
PR 05-OCT-1999; 99WO-US023089.  
PR 29-NOV-1999; 99WO-US028214.  
PR 30-NOV-1999; 99WO-US028313.  
PR 30-NOV-1999; 99WO-US028409.  
PR 01-DEC-1999; 99WO-US028301.  
PR 01-DEC-1999; 99WO-US028634.  
PR 02-DEC-1999; 99WO-US028551.  
PR 02-DEC-1999; 99WO-US028564.  
PR 02-DEC-1999; 99WO-US028565.  
PR 16-DEC-1999; 99WO-US030095.  
PR 20-DEC-1999; 99WO-US030911.  
PR 20-DEC-1999; 99WO-US030999.  
PR 22-DEC-1999; 99WO-US030720.  
PR 30-DEC-1999; 99WO-US031243.  
PR 30-DEC-1999; 99WO-US031274.  
PR 05-JAN-2000; 2000WO-US000219.  
PR 06-JAN-2000; 2000WO-US000277.  
PR 06-JAN-2000; 2000WO-US000376.  
PR 11-FEB-2000; 2000WO-US003565.  
PR 18-FEB-2000; 2000WO-US004341.  
PR 18-FEB-2000; 2000WO-US004342.  
PR 22-FEB-2000; 2000WO-US004414.  
PR 24-FEB-2000; 2000WO-US004914.  
PR 24-FEB-2000; 2000WO-US005004.  
PR 01-MAR-2000; 2000WO-US005601.  
PR 02-MAR-2000; 2000WO-US005746.  
PR 02-MAR-2000; 2000WO-US005841.  
PR 10-MAR-2000; 2000WO-US006319.  
PR 15-MAR-2000; 2000WO-US006884.  
PR 20-MAR-2000; 2000WO-US007377.  
PR 21-MAR-2000; 2000WO-US007532.  
PR 30-MAR-2000; 2000WO-US008439.  
PR 17-MAY-2000; 2000WO-US013705.  
PR 22-MAY-2000; 2000WO-US014042.  
PR 30-MAY-2000; 2000WO-US014941.  
PR 02-JUN-2000; 2000WO-US015264.  
PR 28-JUL-2000; 2000WO-US020710.  
PR 11-AUG-2000; 2000WO-US022031.  
PR 23-AUG-2000; 2000WO-US023522.  
PR 24-AUG-2000; 2000WO-US023328.  
PR 08-NOV-2000; 2000WO-US030952.  
PR 10-NOV-2000; 2000WO-US030873.  
PR 01-DEC-2000; 2000WO-US032678.  
PR 20-DEC-2000; 2000US-00747259.  
PR 20-DEC-2000; 2000WO-US034956.  
PR 28-FEB-2001; 2001US-00796498.  
PR 28-FEB-2001; 2001WO-US006520.  
PR 01-MAR-2001; 2001WO-US006666.  
PR 09-MAR-2001; 2001US-00802706.  
PR 14-MAR-2001; 2001US-00808689.  
PR 22-MAR-2001; 2001US-00818744.  
PR 05-APR-2001; 2001US-00828366.  
PR 10-MAY-2001; 2001US-00854208.  
PR 10-MAY-2001; 2001US-00854280.  
PR 18-MAY-2001; 2001US-00860216.  
PR 25-MAY-2001; 2001US-00866028.  
PR 25-MAY-2001; 2001US-00866034.  
PR 25-MAY-2001; 2001WO-US017092.  
PR 01-JUN-2001; 2001US-00872035.  
PR 01-JUN-2001; 2001WO-US017800.  
PR 05-JUN-2001; 2001US-00874503.  
PR 14-JUN-2001; 2001US-00882636.  
PR 19-JUN-2001; 2001US-00886342.



Db 1261 GCACCAGAAGAGTTCAACCCAGACCGCTTCTCGATGCAGATGGACGTTTCAGAAACA 1320  
Qy 1584 TGAGCGTTCCTGCCCCCTCTCTTAGGAAGCGTGTCTGCTTGGAGAGGGCTTGGCAAA 1643  
Db 1321 TGAGCGTTCCTGCCCCCTCTCTTAGGAAGCGTGTCTGCTTGGAGAGGGCTTGGCAAA 1380  
Qy 1644 AGCGAGCTCTTCTCTCTTTTCAACCACTCTCAAGCCTTCTCCCTGGAGAGCCCGTG 1703  
Db 1381 AGCGAGCTCTTCTCTCTTTTCAACCACTCTCAAGCCTTCTCCCTGGAGAGCCCGTG 1440  
Qy 1704 CCGCGCGACACCTTGAGCCTCAAGCCCAAGCCTCAGTGGCCCTTTTCAACATTTCCCCACAG 1763  
Db 1441 CCGCGCGACACCTTGAGCCTCAAGCCCAAGCCTCAGTGGCCCTTTTCAACATTTCCCCACAG 1500  
Qy 1764 CTTCCAGTGAAGTCCCTCCCACTGACCTTCACTCCACGACGACCCAGATGAAGAA 1823  
Db 1501 CTTCCAGTGAAGTCCCTCCCACTGACCTTCACTCCACGACGACCCAGATGAAGAA 1560  
Qy 1824 GGCAACTTGGAGTGGTGGGTGCCAGGACGGTGCTCCAGCCTCAACAGTGGGATGGA 1883  
Db 1561 GGCAACTTGGAGTGGTGGGTGCCAGGACGGTGCTCCAGCCTCAACAGTGGGATGGA 1620  
Qy 1884 CAGGGTTAATGCTCCAGAGTGTACATGACGGCAGCCACATTTACAGCCTGCAGTTGT 1943  
Db 1621 CAGGGTTAATGCTCCAGAGTGTACATGACGGCAGCCACATTTACAGCCTGCAGTTGT 1680  
Qy 1944 TTTCCGGAGTCTGCCACGCCCCACAGCCTCACTTGACTCATGCTGTGAAGTGCACAA 2003  
Db 1681 TTTCCGGAGTCTGCCACGCCCCACAGCCTCACTTGACTCATGCTGTGAAGTGCACAA 1740  
Qy 2004 CCGCACACCCATACACAACACTACAAGGGCCACAAAGCACTCTGGTTAGCTTTCCACAG 2063  
Db 1741 CCGCACACCCATACACAACACTACAAGGGCCACAAAGCACTCTGGTTAGCTTTCCACAG 1800  
Qy 2064 ACATAAATATAGTCCATCTGCAATCACAGACATAGCCAGTAAACCCACCAACTCCCCCT 2123  
Db 1801 ACATAAATATAGTCCATCTGCAATCACAGACATAGCCAGTAAACCCACCAACTCCCCCT 1860  
Qy 2124 GGATCTGACGCCACAGTGGAGTCTGGCTGTCACTTCAAGGCCACAGAAAGCGGCA 2183  
Db 1861 GGATCTGACGCCACAGTGGAGTCTGGCTGTCACTTCAAGGCCACAGAAAGCGGCA 1920  
Qy 2184 CACATGTTACAGCTTCACAGCCCTCTCCATTCATCGAACTTCTCAGTGTCCCTGTCCCT 2243  
Db 1921 CACATGTTACAGCTTCACAGCCCTCTCCATTCATCGAACTTCTCAGTGTCCCTGTCCCT 1980  
Qy 2244 GGTGCTGGCACAGGAAACAGATGCCCCCTCCGGGGTCAATGCCACCCAGAGACTGTGC 2303  
Db 1981 GGTGCTGGCACAGGAAACAGATGCCCCCTCCGGGGTCAATGCCACCCAGAGACTGTGC 2040  
Qy 2304 TGTCTATGGCCCCAATCATGCTCCCTCTCTTGGCTACACACTCTCCAGCCTGTGACC 2363  
Db 2041 TGTCTATGGCCCCAATCATGCTCCCTCTCTTGGCTACACACTCTCCAGCCTGTGACC 2100  
Qy 2364 ACCGATGTCCACACACCCCAACACATGTTGCCACAGCTACCCACGCTAGCACATCTGCC 2423  
Db 2101 ACCGATGTCCACACACCCCAACACATGTTGCCACAGCTACCCACGCTAACAATCTGCC 2160  
Qy 2424 TGGTCTCCCAAGATATCTTCCCACTGAGACACGCGCCCCCACCAGAGGACAGTCCCCAG 2483  
Db 2161 TGGTCTCCCAAGATATCTTCCCACTGAGACACGCGCCCCCACCAGAGGACAGTCCCCAG 2220  
Qy 2484 CCACCTCTGCAACGAGCCCTGAGTCAACCCCTTTTAAAGCACCTGATTTACCAATG 2543  
Db 2221 CCACCTCTGCAACGAGCCCTGAGTCAACCCCTTTTAAAGCACCTGATTTACCAATG 2280  
Qy 2544 CAACACATCTGGGTCTGGGATATGACACAGAGACTTTTGACATACAGGACCCCTCAGAC 2603  
Db 2281 CAACACATCTGGGTCTGGGATATGACACAGAGACTTTTGACATACAGGACCCCTCAGAC 2340  
Qy 2604 CGAGGAAACCTGCCCAACCCCAACACGCTTATGTAACCCAGTGGAAAGCGGCCCT 2663  
Db 2341 CGAGGAAACCTGCCCAACCCCAACACGCTTATGTAACCCAGTGGAAAGCGGCCCT 2400

Qy 2664 GCTGCCCTCCACACACACATACACTACTGATCTACAGCCCTGTTCCGCGTCAGAG 2723  
Db 2401 GCTGCCCTCCACACACACATACACTACTGATCTACAGCCCTGTTCCGCGTCAGAG 2460  
Qy 2724 TCCCACTAGACCCAGTGGAGGGGTTAGACACCAAGTAGGGCCAGTTTCCAATTCAAC 2783  
Db 2461 TCCCACTAGACCCAGTGGAGGGGTTAGACACCAAGTAGGGCCAGTTTCCAATTCAAC 2520  
Qy 2784 CTGTCAAGGAGTGAAGCCCGATCTGACCTTCTCTTGTGACTTAAGGCTCCGGCTTGGGAATT 2843  
Db 2521 CTGTCAAGGAGTGAAGCCCGATCTGACCTTCTCTTGTGACTTAAGGCTCCGGCTTGGGAATT 2580  
Qy 2844 AAAGTTTGTCTGGCCTTTAGCCTTAAAAAAA 2886  
Db 2581 AAAGTTTGTCTGGCCTTTAGCCTTAAAAAAA 2623

RESULT 13  
ADB19071  
ID ADB19071 standard; cDNA; 2623 BP.  
XX ADB19071;  
XX AC ADB19071;  
XX DT 20-NOV-2003 (first entry)  
XX XX  
DE Novel human secreted and transmembrane protein PRO1906 cDNA.  
XX Human; secreted and transmembrane protein; PRO; gene; ss;  
KW Tumour necrosis factor alpha release; TNF-alpha release;  
KW glucose uptake modulator; FFA uptake modulator;  
KW cell proliferation stimulator; cell differentiation stimulator;  
KW cell differentiation inhibitor; cytokin.  
XX Homo sapiens.  
XX XX  
XX US2003068796-A1.  
XX PD 10-APR-2003.  
XX XX  
XX 15-APR-2002; 2002US-00123261.  
XX XX  
PR 31-MAR-1997; 97WO-US005230.  
PR 12-JUN-1998; 98WO-US012456.  
PR 14-JUL-1998; 98WO-US014552.  
PR 28-AUG-1998; 98WO-US017888.  
PR 10-SEP-1998; 98WO-US018824.  
PR 14-SEP-1998; 98WO-US019093.  
PR 14-SEP-1998; 98WO-US019094.  
PR 14-SEP-1998; 98WO-US019177.  
PR 16-SEP-1998; 98WO-US019330.  
PR 17-SEP-1998; 98WO-US019437.  
PR 07-OCT-1998; 98WO-US021141.  
PR 29-OCT-1998; 98WO-US022991.  
PR 29-OCT-1998; 98WO-US022992.  
PR 20-NOV-1998; 98WO-US024855.  
PR 01-DEC-1998; 98WO-US025108.  
PR 05-JAN-1999; 99WO-US000106.  
PR 08-MAR-1999; 99WO-US005028.  
PR 10-MAR-1999; 99WO-US005190.  
PR 20-APR-1999; 99WO-US008615.  
PR 14-MAY-1999; 99WO-US010733.  
PR 02-JUN-1999; 99WO-US012252.  
PR 01-SEP-1999; 99WO-US020111.  
PR 08-SEP-1999; 99WO-US020594.  
PR 13-SEP-1999; 99WO-US020944.  
PR 15-SEP-1999; 99WO-US021090.  
PR 15-SEP-1999; 99WO-US021547.  
PR 05-OCT-1999; 99WO-US023089.  
PR 29-NOV-1999; 99WO-US028214.  
PR 30-NOV-1999; 99WO-US028313.  
PR 30-NOV-1999; 99WO-US028409.  
PR 01-DEC-1999; 99WO-US028301.





QY 1044 CACAGTCCGGAGGTGACAGCAGCAGGGAACCTGGATGCTTCCGGCCCCGACGTGA 1103  
Db 781 CACAGTCCGGAGGTGACAGCAGCAGGGAACCTGGATGCTTCCGGCCCCGACGTGA 840  
QY 1104 CTTGTGATGCTTCTGCTGGAAGATGCGACAGGAGGAACAAACCCAGGACAGAAAT 1163  
Db 841 CTTGTGATGCTTCTGCTGGAAGATGCGACAGGAGGAACAAACCCAGGACAGAAAT 900  
QY 1164 CACCAACAGACATGCTGATGACAGTCAATTTATTTGCTGTTTGTGGGACGATGACGGT 1223  
Db 901 CACCAACAGACATGCTGATGACAGTCAATTTATTTGCTGTTTGTGGGACGATGACGGT 960  
QY 1224 CAGCACACAGGTCCGCTATACCTCTGCTGATGAAATACCTCATGTCCCAAAAGTG 1283  
Db 961 CAGCACACAGGTCCGCTATACCTCTGCTGATGAAATACCTCATGTCCCAAAAGTG 1020  
QY 1284 GGTACGTGAGAGTGAATCGGAGCTGGGGCTGGCCAGGACCAAGCCTTAGGGACCG 1343  
Db 1021 GGTACGTGAGAGTGAATCGGAGCTGGGGCTGGCCAGGACCAAGCCTTAGGGACCG 1080  
QY 1344 TACCCGCTCCCTTACACGAGCGGTTCTGATGAGCGGAGCGGCTGCTGGCGTGGT 1403  
Db 1081 TACCCGCTCCCTTACACGAGCGGTTCTGATGAGCGGAGCGGCTGCTGGCGTGGT 1140  
QY 1404 GCCCATGGGAATACCCCGCACCCCTCATGCGGACCAACCGCTTCCGAGGGTACACCTGCC 1463  
Db 1141 GCCCATGGGAATACCCCGCACCCCTCATGCGGACCAACCGCTTCCGAGGGTACACCTGCC 1200  
QY 1464 CAGGGCA CGGAGTCTTCCCTCTTGGCTTCATCTGATGACCCCAACATCTTCAA 1523  
Db 1201 CAGGGCA CGGAGTCTTCCCTCTTGGCTTCATCTGATGACCCCAACATCTTCAA 1260  
QY 1524 GCACCCAGAGAGTTCAACCCAGACCGTTTCTGGATGCGATGGAAGCGTTTCAGGAAGCA 1583  
Db 1261 GCACCCAGAGAGTTCAACCCAGACCGTTTCTGGATGCGATGGAAGCGTTTCAGGAAGCA 1320  
QY 1584 TGAGGGCTTCTGCTTCTTAGGGAAGCGTGTCTGCTTGGAGAGCGCTTGGCAA 1643  
Db 1321 TGAGGGCTTCTGCTTCTTAGGGAAGCGTGTCTGCTTGGAGAGCGCTTGGCAA 1380  
QY 1644 AGCGAGCTTCTGCTTCTTACCAACCATCTTACAGCGCTTCTCCCTGGAGAGCGCGTG 1703  
Db 1381 AGCGAGCTTCTGCTTCTTACCAACCATCTTACAGCGCTTCTCCCTGGAGAGCGCGTG 1440  
QY 1704 CCGCGCGACACCTGAGCTCAAGCCACCGTCAAGTGGCTTTTCAAATTCCTCCCGACG 1763  
Db 1441 CCGCGCGACACCTGAGCTCAAGCCACCGTCAAGTGGCTTTTCAAATTCCTCCCGACG 1500  
QY 1764 CTTCCAGCTGCAAGTCCGTCCTTCCACTGACCTTCACTCCACACGACAGACAGATGAGGAA 1823  
Db 1501 CTTCCAGCTGCAAGTCCGTCCTTCCACTGACCTTCACTCCACACGACAGATGAGGAA 1560  
QY 1824 GGCACCTTGAAGTGGTGGTGCCAGGACGGTGCCTCCAGCCTCAACAGTGGGATGGA 1883  
Db 1561 GGCACCTTGAAGTGGTGGTGCCAGGACGGTGCCTCCAGCCTCAACAGTGGGATGGA 1620  
QY 1884 CAGGGTTAATGCTCCAGAGTGTACACTGAGGAGGACCAATTTACAGCCTGCAATTTGT 1943  
Db 1621 CAGGGTTAATGCTCCAGAGTGTACACTGAGGAGGACCAATTTACAGCCTGCAATTTGT 1680  
QY 1944 TTTCCGAGTCTGCTCCAGCGGCCACACGCTCACTTGACTCATGCTGCTAAGATGACAA 2003  
Db 1681 TTTCCGAGTCTGCTCCAGCGGCCACACGCTCACTTGACTCATGCTGCTAAGATGACAA 1740  
QY 2004 CGCACACCCATACAACTAAGGGGCCAAAGCAACTGCTGGGTAGCTTTCCACAG 2063  
Db 1741 CGCACACCCATACAACTAAGGGGCCAAAGCAACTGCTGGGTAGCTTTCCACAG 1800  
QY 2064 ACATAAATAGTCCATCTGCATACAGACCATAGCCAGGTAACCCCACTCCCT 2123  
Db 1801 ACATAAATAGTCCATCTGCATACAGACCATAGCCAGGTAACCCCACTCCCT 1860

QY 2124 GGATCTGAGCCACACGTTGGAGTCTGGCTGTCACTTTCAAGCCACAGAAACGGGCA 2183  
Db 1861 GGATCTGAGCCACACGTTGGAGTCTGGCTGTCACTTTCAAGCCACAGAAACGGGCA 1920  
QY 2184 CACATGTTTACAGCTTCAACGCCCTCTCATTTCAATCGAACTTCTCAGTGTCCCTGCTCCCT 2243  
Db 1921 CACATGTTTACAGCTTCAACGCCCTCTCATTTCAATCGAACTTCTCAGTGTCCCTGCTCCCT 1980  
QY 2244 GGTGCTGGCACAGGGAACAGCATGCCCTTCCGGGGTCAATGCCACCCAGAGACTGTGCG 2303  
Db 1981 GGTGCTGGCACAGGGAACAGCATGCCCTTCCGGGGTCAATGCCACCCAGAGACTGTGCG 2040  
QY 2304 TGTCTATGGCCCCAACTCATGCTCCCTCTCTTGGGTACACACTCTTCCAGCCTGTGACC 2363  
Db 2041 TGTCTATGGCCCCAACTCATGCTCCCTCTCTTGGGTACACACTCTTCCAGCCTGTGACC 2100  
QY 2364 ACCGATGTCCACACACACCCCAACCACTTGTCCACACAGCTACCCACGTAACCAATCGTCC 2423  
Db 2101 ACCGATGTCCACACACACCCCAACCACTTGTCCACACAGCTACCCACGTAACCAATCGTCC 2160  
QY 2424 TGGCTCCCGAGAGTATCTTCCACTGAGACACGCGCCGCCCCACACAGAGGCAAGTCCCCAG 2483  
Db 2161 TGGCTCCCGAGAGTATCTTCCACTGAGACACGCGCGCCGCCCCACACAGAGGCAAGTCCCCAG 2220  
QY 2484 CCACCTCTGCAACTGAGCCTCAGTCAACCCCTTTTAAAGACCCCTGATTTTACCAATG 2543  
Db 2221 CCACCTCTGCAACTGAGCCTCAGTCAACCCCTTTTAAAGACCCCTGATTTTACCAATG 2280  
QY 2544 CAAACACATCTGGTCTGCGATTATGACAGAGACTTTGGACATACAGGACCCCTGAGAC 2603  
Db 2281 CAAACACATCTGGTCTGCGATTATGACAGAGACTTTGGACATACAGGACCCCTGAGAC 2340  
QY 2604 CGGAGGAACACTGCCCAACCCCAACACGCTTATGTAACCAAGCGGCGCCCT 2663  
Db 2341 CGGAGGAACACTGCCCAACCCCAACACGCTTATGTAACCAAGCGGCGCCCT 2400  
QY 2664 GCTGCCCTCCACACACATACACACTGATGATCTAGACCCCTGTTGGGCTGAGAG 2723  
Db 2401 GCTGCCCTCCACACACATACACACTGATGATCTAGACCCCTGTTGGGCTGAGAG 2460  
QY 2724 TCCCACTAGACCCAGTGGAGGGTTAGAGACCAAGTAGGGCCAGTTTCCCAATTCACC 2783  
Db 2461 TCCCACTAGACCCAGTGGAGGGTTAGAGACCAAGTAGGGCCAGTTTCCCAATTCACC 2520  
QY 2784 CTCTCAGGGAGTGGAGCGGATCTGAGCTTCTTGTGACTTAAAGGTCGGCTTGGGAAT 2843  
Db 2521 CTCTCAGGGAGTGGAGCGGATCTGAGCTTCTTGTGACTTAAAGGTCGGCTTGGGAAT 2580  
QY 2844 AAAGTTGTTTCTGGCTTTAGCTTAAAGGTCGGCTTGGGAAT 2886  
Db 2581 AAAGTTGTTTCTGGCTTTAGCTTAAAGGTCGGCTTGGGAAT 2623

## RESULT 14

ADB27612

ID ADB27612 standard; cDNA; 2623 BP.

XX

AC ADB27612;

XX

DT 20-NOV-2003 (first entry)

XX

DE cDNA encoding human PRO polypeptide #32.

XX

Human; gene; ss; PRO; secreted polypeptide; transmembrane polypeptide;  
tumour necrosis factor-alpha; TNF-alpha; chondrocyte cell; tumour;  
cancer; adrenal; lung; colon; breast; prostate; kidney; cervix;  
liver; microvascular endothelial cell; glucose; FFA;  
skeletal muscle cell; adipocyte cell; pericyte cell;  
inner ear utricular supporting cell; T-lymphocyte cell;  
endothelial cell tube formation; bone disorder; cartilage disorder;  
sports injury; proteoglycan; articular cartilage defect; osteoarthritis;  
rheumatoid arthritis; haemoglobin-associated disorder thalassemia;  
immune system cell infiltration.



QY 1464 CCAGGCGACGAGGCTTCTCCCTCTCTGGCTCCATCTCTGCATGACCCCAACATCTTCAA 1523  
Db 1201 CCAGGCGACGAGGCTTCTCCCTCTCTGGCTCCATCTCTGCATGACCCCAACATCTTCAA 1260  
QY 1524 GCACCCAGAGAGTTCAACCCAGACCGTTTCTGGATGAGATGAGCGTTTTCAGGAAGCA 1583  
Db 1261 GCACCCAGAGAGTTCAACCCAGACCGTTTCTGGATGAGATGAGCGTTTTCAGGAAGCA 1320  
QY 1584 TGAGCGCTTCTCCCTCTCTCTCTAGGAGAGCGTGTCTGCTTGGAGAGGGCTTGCAAA 1643  
Db 1321 TGAGCGCTTCTCCCTCTCTCTCTAGGAGAGCGTGTCTGCTTGGAGAGGGCTTGCAAA 1380  
QY 1644 AGCGAGCTTCTCTCTCTCTTCCACCAACATCTCAAGCCACCGTCAAGTGGCCCTTTCAACATTTCCCCACG 1763  
Db 1381 AGCGAGCTTCTCTCTCTCTTCCACCAACATCTCAAGCCACCGTCAAGTGGCCCTTTCAACATTTCCCCACG 1440  
QY 1704 CCGCGCGAGACCTTGAGCTCAAGCCACCGTCAAGTGGCCCTTTCAACATTTCCCCACG 1500  
Db 1441 CCGCGCGAGACCTTGAGCTCAAGCCACCGTCAAGTGGCCCTTTCAACATTTCCCCACG 1560  
QY 1764 CTTCCAGCTGCAAGTCCGCTCCCACTGACCTTCACTCCACCGGAGCAGATGAAGAA 1823  
Db 1501 CTTCCAGCTGCAAGTCCGCTCCCACTGACCTTCACTCCACCGGAGCAGATGAAGAA 1560  
QY 1824 GGCAACTTGGAGTGGTGGGTCGCCAGGACCGTCCAGCTCAACAGTGGGATGGA 1883  
Db 1561 GGCAACTTGGAGTGGTGGGTCGCCAGGACCGTCCAGCTCAACAGTGGGATGGA 1620  
QY 1884 CAGGGTAAATGTTCTCCAGAGTGTACATGCGAGGAGCCACATTTACAGCCCTGCAAGTGT 1943  
Db 1621 CAGGGTAAATGTTCTCCAGAGTGTACATGCGAGGAGCCACATTTACAGCCCTGCAAGTGT 1680  
QY 1944 TTTCCGAGTCTGTCCAGGSCCCACAGCTCACTGACTCATGCTCATGCTTAAGATGCACAA 2003  
Db 1681 TTTCCGAGTCTGTCCAGGSCCCACAGCTCACTGACTCATGCTCATGCTTAAGATGCACAA 1740  
QY 2004 CCGCACACCCATACACAACTACAAGGSCCAACAAAGCAACTCTGGTGTAGCTTTCCACAG 2063  
Db 1741 CCGCACACCCATACACAACTACAAGGSCCAACAAAGCAACTCTGGTGTAGCTTTCCACAG 1800  
QY 2064 ACATAATATAGTCCATCTGCAATCAAGCATAGCAGGTAACCCACCACTCCCT 2123  
Db 1801 ACATAATATAGTCCATCTGCAATCAAGCATAGCAGGTAACCCACCACTCCCT 1860  
QY 2124 GGATCTGACGACACAGTGGAGTCTGGTGTACCTTCAAGCATAGCAGGTAACCCACCACTCCCT 2183  
Db 1861 GGATCTGACGACACAGTGGAGTCTGGTGTACCTTCAAGCATAGCAGGTAACCCACCACTCCCT 1920  
QY 2184 CACATGTTCAAGTCAACAGCCCTCTCCATTTCAATCGAACTTCTCAGTGTCCCTGTCCCT 2243  
Db 1921 CACATGTTCAAGTCAACAGCCCTCTCCATTTCAATCGAACTTCTCAGTGTCCCTGTCCCT 1980  
QY 2244 GGTGCTTGGCACAGGGAACAGATGCCCCCTCGGGGTCAATGCAACCCAGAGACTGTGCG 2303  
Db 1981 GGTGCTTGGCACAGGGAACAGATGCCCCCTCGGGGTCAATGCAACCCAGAGACTGTGCG 2040  
QY 2304 TGTCTATGGCCCAACTCATGCTCCCTCTCTTGGCTACACCACTCTCCAGAGCTGTGACC 2363  
Db 2041 TGTCTATGGCCCAACTCATGCTCCCTCTCTTGGCTACACCACTCTCCAGAGCTGTGACC 2100  
QY 2364 ACCGATGTCCACACACCCCAACCACTTGTCCACACAGCTTACCCAGTACGACATCGTCC 2423  
Db 2101 ACCGATGTCCACACACCCCAACCACTTGTCCACACAGCTTACCCAGTACGACATCGTCC 2160  
QY 2424 TGGCTCCCGAGAGTATCTCCGACTGAGACACCGCCGCCCCCAACAGAGGACAGTCCCCCAG 2483  
Db 2161 TGGCTCCCGAGAGTATCTCCGACTGAGACACCGCCGCCCCCAACAGAGGACAGTCCCCCAG 2220  
QY 2484 CCACCTCTGCAACTGAGGCTCAGTCAACCCCTTTTAAAGCACCCTGATTTACCAAAATG 2543  
Db 2221 CCACCTCTGCAACTGAGGCTCAGTCAACCCCTTTTAAAGCACCCTGATTTACCAAAATG 2280

QY 2544 CAAACACATCTGGTCTGGATTTATGCAACAGAGACTTTGGACATACGAGACCCCTCAGAC 2603  
Db 2281 CAAACACATCTGGTCTGGATTTATGCAACAGAGACTTTGGACATACGAGACCCCTCAGAC 2340  
QY 2604 CGGAGGACACCTGCGCCACCCCAACAGCGTCTTATGTAACCAAGTGGAAAGCGGCCCT 2663  
Db 2341 CGGAGGACACCTGCGCCACCCCAACAGCGTCTTATGTAACCAAGTGGAAAGCGGCCCT 2400  
QY 2664 GCTGCCCTTCCACACACACATACACACTCAGTCTACAGCCCTGTTGGGCTCAGAG 2723  
Db 2401 GCTGCCCTTCCACACACATACACACTCAGTCTACAGCCCTGTTGGGCTCAGAG 2460  
QY 2724 TCCCACTAGAGCCAGTGAAGGGTTAGAGACCAAGTAGGGGCCAGTTTCCAATTCAAC 2783  
Db 2461 TCCCACTAGAGCCAGTGAAGGGTTAGAGACCAAGTAGGGGCCAGTTTCCAATTCAAC 2520  
QY 2784 CTGTGAGGAGTGAGCCGAGTCTGAGCTTCCCTTGTGACTTAAGGTCGGCTTGGGAAT 2843  
Db 2521 CTGTGAGGAGTGAGCCGAGTCTGAGCTTCCCTTGTGACTTAAGGTCGGCTTGGGAAT 2580  
QY 2844 AAAGTTTCTTCTGGCTTTAGCCCTTAAATAAAAAAAAAAAAAA 2886  
Db 2581 AAAGTTTCTTCTGGCTTTAGCCCTTAAATAAAAAAAAAAAAAA 2623

RESULT 15  
ADAB6091  
ID ADAB6091 standard; cDNA; 2623 BP.  
XX  
AC ADAB6091;  
XX  
DT 20-NOV-2003 (first entry)  
XX  
DE Novel human secreted and transmembrane protein PRO1906 cDNA.  
XX  
KW Human; secreted and transmembrane protein; PRO; gene; ss;  
KW Tumour necrosis factor alpha release; TNF-alpha release;  
KW Glucose uptake modulator; FFA uptake modulator;  
KW cell proliferation stimulator; cell differentiation stimulator;  
KW cell differentiation inhibitor; cytokine release stimulator;  
KW lung tumour; colon tumour; breast tumour; prostate tumour; rectal tumour;  
KW cervical tumour; liver tumour; chromosome mapping; gene mapping;  
KW gene therapy; chromosome identification; chromosome marker.  
XX  
OS Homo sapiens.  
XX  
FN US2003082711-A1.  
XX  
PD 01-MAY-2003.  
XX  
PF 16-MAY-2002; 2002US-00147508.  
XX  
PR 02-JUL-1998; 98US-0091519P.  
PR 02-JUN-1999; 99WO-US012252.  
PR 07-JUL-1999; 99US-0143048P.  
PR 25-AUG-1999; 99US-00380137.  
PR 30-MAR-2000; 2000WO-US008439.  
PR 01-DEC-2000; 2000WO-US032678.  
PR 19-DEC-2001; 2001US-00028072.  
XX  
(GETH ) GENENTECH INC.  
XX  
XX Baker KP, Beresini M, Deforge L, Desnoyers L, Filvaroff E, Gao W;  
PI Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;  
PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;  
XX WPI; 2003-786914/74.  
DR P-PSDB; ADAB6092.  
XX  
PT New PRO nucleic acid, useful for preparing a composition for treating  
PT e.g., tumor or for tissue typing.  
XX  
PS Claim 2; Fig 63; 637pp; English.

XX The invention describes 305 nucleic acids encoding PRO (secreted and  
CC transmembrane) polypeptides (I). (I) is useful for stimulating the  
CC release of TNF- $\alpha$  from human blood, for modulating the uptake of  
CC glucose or FFA by skeletal muscle cells or adipocyte cells, for  
CC stimulating the proliferation or differentiation of chondrocyte cells,  
CC for stimulating the proliferation of or gene expression in pericyte  
CC cells, for stimulating the release of proteoglycans from cartilage, for  
CC stimulating the proliferation of inner ear utricular supporting cells,  
CC for stimulating the proliferation of T-lymphocyte cells, for stimulating  
CC the release of a cytokine from PMBC cells, for inhibiting the binding of  
CC A-peptide to factor VIIa, for inhibiting the differentiation of adipocyte  
CC cells, for stimulating proliferation of endothelial cells, for detecting  
CC the presence of tumour in a mammal. The tumour is lung, colon, breast,  
CC prostate, rectal, cervical or liver tumour. The oligonucleotide probes  
CC are useful for isolating genomic and cDNA nucleotide sequences or  
CC antisense probes. (I) is also useful as therapeutic agent. PRO is useful  
CC in assays to identify other proteins or molecules involved in binding  
CC interaction. A polynucleotide (II) encoding (I) is useful in chromosome  
CC and gene mapping, in generation of antisense RNA and DNA, in the  
CC preparation of PRO polypeptide, for generating transgenic animals or  
CC knockout animals which in turn are useful in the development and  
CC screening of therapeutically useful reagents, in gene therapy, for  
CC chromosome identification, as chromosome marker, and for generating  
CC probes. An anti-(I)-antibody is useful in diagnostic assays for PRO, e.g.  
CC detecting its expression in specific cells, tissues or serum, and for  
CC affinity purification of PRO from recombinant cell culture or natural  
CC sources. (I) and (II) are useful for tissue typing. This sequence encodes  
CC a novel human secreted and transmembrane PRO polypeptide.

XX  
SQ Sequence 2623 BP; 552 A; 865 C; 704 G; 502 T; 0 U; 0 Other;

Query Match 89.0%; Score 2621.4; DB 8; Length 2623;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2622; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	264	GGGAGCGCTGGGAGGAGGAGGAGCGACCTGCGGAGATGAGGCGACCGGACCTG 323
DB	1	GGGAGCGCTGGGAGGAGGAGGAGCGACCTGCGGAGATGAGGCGACCGGACCTG 60
QY	324	GGCGCTGCTGCTGGCGCTGGCGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 383
DB	61	GGCGCTGCTGCTGGCGCTGGCGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
QY	384	GGCGCGAGGCGCACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 443
DB	121	GGCGCGAGGCGCACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180
QY	444	GCTACGCGCGCGCGCGCGCTGATTACGCGGCTCATGCGGCTGAGTAAGAGTACG 503
DB	181	GCTACGCGCGCGCGCGCGCTGATTACGCGGCTCATGCGGCTGAGTAAGAGTACG 240
QY	504	GTTACCATCTACTGCGGACCTTGGCGGCTGTGGTGTCTGTTGGGAGGAGGCTGT 563
DB	241	GTTACCATCTACTGCGGACCTTGGCGGCTGTGGTGTCTGTTGGGAGGAGGCTGT 300
QY	564	GGGAGGCGCGCTGGGAGGTGAGGTGAGGTGAGGTGAGGTGAGGTGAGGTGAGGTG 623
DB	301	GGGAGGCGCGCTGGGAGGTGAGGTGAGGTGAGGTGAGGTGAGGTGAGGTGAGGTG 360
QY	624	GGAAGGACCTTTGATGCGCATGGGCTTTCTCTCCAAAGCGGAGCGGTGGAGGAG 683
DB	361	GGAAGGACCTTTGATGCGCATGGGCTTTCTCTCCAAAGCGGAGCGGTGGAGGAG 420
QY	684	GAGGAAGTTTACCATGCTGCTGCGGAGCTTGGGAGTGGGAGTGGGAGTGGGAG 743
DB	421	GAGGAAGTTTACCATGCTGCTGCGGAGCTTGGGAGTGGGAGTGGGAGTGGGAG 480
QY	744	GCTGATCGAGCGAGGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 803
DB	481	GCTGATCGAGCGAGGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
QY	804	ATTGATCCCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 863

DB	541	ATTGATCCCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
QY	864	TGGCGCTCGCTTCTCCCTATGAGGATAGGAGTTCAGGCGCGTGGTCCGGGAGCTGG 923
DB	601	TGGCGCTCGCTTCTCCCTATGAGGATAGGAGTTCAGGCGCGTGGTCCGGGAGCTGG 660
QY	924	TACCTGCTGGAGTCACTCCAGGGGGGTGAGACCTACGAGATGTTCTCTCGTGTCT 983
DB	661	TACCTGCTGGAGTCACTCCAGGGGGGTGAGACCTACGAGATGTTCTCTCGTGTCT 720
QY	984	GGGCGCGCTGGCAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 1043
DB	721	GGGCGCGCGCTGGCAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 780
QY	1044	CACAGTCGGGAGGTGAGGAGCGACCGAGGAACTGATGCTTCGGGCGCGCGCGCTGA 1103
DB	781	CACAGTCGGGAGGTGAGGAGCGACCGAGGAACTGATGCTTCGGGCGCGCGCGCTGA 840
QY	1104	CCTTGTGATGCTTCTCTGCTGAGATGCGACAGGAGGAAAGAAACCGAGGACAGAT 1163
DB	841	CCTTGTGATGCTTCTCTGCTGAGATGCGACAGGAGGAAAGAAACCGAGGACAGAT 900
QY	1164	CACCAACAGAACATGCTGATGACAGTCAATTTATTTGCTGTTGCTGGGACGATGCG 1223
DB	901	CACCAACAGAACATGCTGATGACAGTCAATTTATTTGCTGTTGCTGGGACGATGCG 960
QY	1224	CAGCACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1283
DB	961	CAGCACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
QY	1284	GGTATGAGGAGTGAATCGGGAGCTGGGGCTGGCGGAGCGACCGAGGAGGAGCGG 1343
DB	1021	GGTATGAGGAGTGAATCGGGAGCTGGGGCTGGCGGAGCGACCGAGGAGGAGCGG 1080
QY	1344	TACCGCGCTTCTTACCGAGCGGCTTCTGATGAGGCGGAGCGGCTGCTGCGCTGG 1403
DB	1081	TACCGCGCTTCTTACCGAGCGGCTTCTGATGAGGCGGAGCGGCTGCTGCGCTGG 1140
QY	1404	GCCATGGGAATACCCCGCACCTCATGCGGAGCGACCGCTTCCGAGGAGTACAC 1463
DB	1141	GCCATGGGAATACCCCGCACCTCATGCGGAGCGACCGCTTCCGAGGAGTACAC 1200
QY	1464	CCAGGCGAGGAGGCTTCTTCCCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1523
DB	1201	CCAGGCGAGGAGGCTTCTTCCCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
QY	1524	GCACCCAGAGAGTTCACCCAGACCGCTTCTGAGTGCAGATGAGCGGTTCAAGGAG 1583
DB	1261	GCA CCAGAGAGTTCACCCAGACCGCTTCTGAGTGCAGATGAGCGGTTCAAGGAG 1320
QY	1584	TGAGGCGTTCCTGCGCTTCTTCTTGGGAGCGGCTGCTGCTTGGAGGCGGCTGGCAA 1643
DB	1321	TGAGGCGTTCCTGCGCTTCTTCTTGGGAGCGGCTGCTGCTTGGAGGCGGCTGGCAA 1380
QY	1644	AGCGGAGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1703
DB	1381	AGCGGAGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1440
QY	1704	CCCGCGGAGCAGCTGAGCGCTCAAGCCCGCTGAGTGGCGCTTTCACATTTCCCGCAG 1763
DB	1441	CCCGCGGAGCAGCTGAGCGCTCAAGCCCGCTGAGTGGCGCTTTCACATTTCCCGCAG 1500
QY	1764	CTTCCAGTGCAGTCCGCTCCCACTGACCTTCACTCCAGCGAGCCAGATGAGGAA 1823
DB	1501	CTTCCAGTGCAGTCCGCTCCCACTGACCTTCACTCCAGCGAGCCAGATGAGGAA 1560
QY	1824	GGCAACTGGAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1883
DB	1561	GGCAACTGGAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1620
QY	1884	CAGGCTTAAGTCTCCAGAGTGTACCTGAGGAGCGGAGCAATTTACAGCGCTGCA 1943

Db 1621 CAGGGTTAATGTCTCCAGAGTGTACACTGCGGCGAGCCACATTTTACACGCGCTGAGTTGT 1680  
QY 1944 TTTCGGAGTGTGTCCCGACCGGCCACAGCTCAGCTGACTGATGTCTGATGATGACAA 2003  
Db 1681 TTTCGGAGTGTGTCCCGACCGGCCACAGCTCAGCTGACTGATGTCTGATGATGACAA 1740  
QY 2004 CCGCACACCCCATACACAACTTACAAGGGCCACAAAGCAACTGCTGGGTTAGCTTTCCACAG 2063  
Db 1741 CCGCACACCCCATACACAACTTACAAGGGCCACAAAGCAACTGCTGGGTTAGCTTTCCACAG 1800  
QY 2064 ACATAAATATAGTCCATCTGCAATCACAAGCACATAGCCAGGTAAACCCACCAACTCCCT 2123  
Db 1801 ACATAAATATAGTCCATCTGCAATCACAAGCACATAGCCAGGTAAACCCACCAACTCCCT 1860  
QY 2124 GGATCTGAGCCCGACACAGTGGAGTCTGGCTGTACCTTCAAGCCACAGAAACGGCCA 2183  
Db 1861 GGATCTGAGCCCGACACAGTGGAGTCTGGCTGTACCTTCAAGCCACAGAAACGGCCA 1920  
QY 2184 CACATGTTACAGCTCACACGCCCTCTCCATTCATCGAACTTCTCAGTGTCCCTGTCCCT 2243  
Db 1921 CACATGTTACAGCTCACACGCCCTCTCCATTCATCGAACTTCTCAGTGTCCCTGTCCCT 1980  
QY 2244 GGTGCTGGGACAGGGAACAGCATGCCCCCTCCGGGTTCATGCCACCCAGAGACTGTGC 2303  
Db 1981 GGTGCTGGGACAGGGAACAGCATGCCCCCTCCGGGTTCATGCCACCCAGAGACTGTGC 2040  
QY 2304 TGTCTATGGCCCCAACTCATGCTCCCTCTCTGGCTACACACTCTCCAGCCTGTGACC 2363  
Db 2041 TGTCTATGGCCCCAACTCATGCTCCCTCTCTGGCTACACACTCTCCAGCCTGTGACC 2100  
QY 2364 ACCGATGTCCACACACCCCAACACATTTGTCCACACAGCTACCCACGTAACGACATCGTCC 2423  
Db 2101 ACCGATGTCCACACACCCCAACACATTTGTCCACACAGCTACCCACGTAACGACATCGTCC 2160  
QY 2424 TGGCTCCCGAGAGTATCTTCCACTGAGACACGCGCCCCCACAGAGGACAGTCCCCAG 2483  
Db 2161 TGGCTCCCGAGAGTATCTTCCACTGAGACACGCGCCCCCACAGAGGACAGTCCCCAG 2220  
QY 2484 CCACCTCTGCAACTGACGCTCAGTCAACCCCTTTTAAAGCACCTGATTTACCAAATG 2543  
Db 2221 CCACCTCTGCAACTGACGCTCAGTCAACCCCTTTTAAAGCACCTGATTTACCAAATG 2280  
QY 2544 CAAACACATCTGGGTCTGCGATTATGCACAGAGACTTTGGACATACGAGGACCCCTCAGAC 2603  
Db 2281 CAAACACATCTGGGTCTGCGATTATGCACAGAGACTTTGGACATACGAGGACCCCTCAGAC 2340  
QY 2604 CGGAGGACACCTGCCCAACCCCAACACGCTGCTTATGTAAACGTTGGAAGCGGCCCT 2663  
Db 2341 CGGAGGACACCTGCCCAACCCCAACACGCTGCTTATGTAAACGTTGGAAGCGGCCCT 2400  
QY 2664 GTGCCCCCTCCACACACATACACTCACTGATCTACAGCCCCCTGTTCCGGGTGAGAG 2723  
Db 2401 GTGCCCCCTCCACACACATACACTCACTGATCTACAGCCCCCTGTTCCGGGTGAGAG 2460  
QY 2724 TCCCCACTAGACCCAGTGAAGGGTTAGAGACCAAGTAGGGCCAGTTTCCCAATTCAAC 2783  
Db 2461 TCCCCACTAGACCCAGTGAAGGGTTAGAGACCAAGTAGGGCCAGTTTCCCAATTCAAC 2520  
QY 2784 CTGTGAGGGAGTGAAGGATCTGAGTTCCTTGTGACTTAAAGGTCGGCTTGGGAAT 2843  
Db 2521 CTGTGAGGGAGTGAAGGATCTGAGTTCCTTGTGACTTAAAGGTCGGCTTGGGAAT 2580  
QY 2844 AAAGTTTGTTCCTGGCTTTAGCCTAAAAAATAAAAAA 2886  
Db 2581 AAAGTTTGTTCCTGGCTTTAGCCTAAAAAATAAAAAA 2623

Search completed: September 17, 2004, 08:41:28  
Job time : 1094.48 secs

**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 17, 2004, 07:38:50 : Search time 188,673 Seconds

(without alignments)  
8659.273 Million cell updates/sec

Title: US-10-669-693-1

Perfect score: 2944

Sequence: 1 tttcttctgttcttactc.....aaaaaaaaaaaaaaaaaaaaa 2944

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA:\*

1: /cgn2\_6/ptodata/2/ina/5A COMB.seq:\*

2: /cgn2\_6/ptodata/2/ina/5B COMB.seq:\*

3: /cgn2\_6/ptodata/2/ina/6A COMB.seq:\*

4: /cgn2\_6/ptodata/2/ina/6B COMB.seq:\*

5: /cgn2\_6/ptodata/2/ina/PCBUS COMB.seq:\*

6: /cgn2\_6/ptodata/2/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	560.8	19.0	567	4	US-09-023-655-694
2	475.2	16.1	1485	4	US-08-277-031B-8
3	475.2	16.1	1485	4	US-08-277-031B-9
4	473.2	16.1	1740	2	US-08-750-703-1
5	469.2	15.9	2907	4	US-09-023-655-1053
6	462.8	15.7	1476	4	US-08-277-031B-10
7	461.2	15.7	1737	2	US-08-750-703-2
8	429.4	14.6	1825	4	US-09-023-655-1061
9	328	11.1	341	4	US-09-023-655-459
10	324	11.0	1473	4	US-08-277-031B-15
11	322.4	11.0	1854	1	US-08-201-118-4
12	322.4	11.0	1854	2	US-08-238-821B-4
13	322.4	11.0	1854	4	US-09-023-655-1056
14	322.4	11.0	1854	5	PCT-US95-05744-4
15	321.8	10.9	1419	2	US-08-194-981E-4
16	321.6	10.9	1892	1	US-08-201-118-14
17	321.6	10.9	1892	2	US-08-238-821B-14
18	321.6	10.9	1892	5	PCT-US95-05744-14
19	320.8	10.9	1473	4	US-08-277-031B-2
20	320.8	10.9	1591	2	US-08-194-981E-3
21	320.8	10.9	1845	1	US-09-023-655-1057
22	320.8	10.9	1852	4	US-08-201-118-10
23	320.8	10.9	1852	2	US-08-238-821B-10
24	320.8	10.9	1852	5	PCT-US95-05744-10
25	319.2	10.8	1746	1	US-08-201-118-2
26	319.2	10.8	1746	2	US-08-238-821B-2
27	319.2	10.8	1746	4	US-09-023-655-1059

28 319.2 10.8 1746 5 PCT-US95-05744-2 Sequence 2, Appli  
29 318.2 10.8 1482 2 US-08-194-981E-1 Sequence 1, Appli  
30 318 10.8 1482 4 US-08-277-031B-3 Sequence 3, Appli  
31 306.2 10.4 1419 2 US-08-194-981E-2 Sequence 2, Appli  
32 296.8 10.1 1473 4 US-08-277-031B-14 Sequence 14, Appli  
33 296.8 10.1 1995 4 US-09-023-655-1055 Sequence 1055, Ap  
34 296.8 10.1 2009 1 US-08-201-118-6 Sequence 6, Appli  
35 296.8 10.1 2009 2 US-08-238-821B-6 Sequence 6, Appli  
36 296.8 10.1 2009 5 PCT-US95-05744-6 Sequence 6, Appli  
37 295.2 10.0 2258 1 US-08-201-118-12 Sequence 12, Appli  
38 295.2 10.0 2258 2 US-08-238-821B-12 Sequence 12, Appli  
39 295.2 10.0 2258 4 US-09-023-655-995 Sequence 995, App  
40 295.2 10.0 2258 5 PCT-US95-05744-12 Sequence 12, Appli  
41 288.4 9.8 1473 4 US-08-277-031B-11 Sequence 11, Appli  
42 285.2 9.7 1473 4 US-08-277-031B-12 Sequence 12, Appli  
43 285.2 9.7 1829 1 US-08-201-118-8 Sequence 8, Appli  
44 285.2 9.7 1829 2 US-08-238-821B-8 Sequence 8, Appli  
45 285.2 9.7 1829 5 PCT-US95-05744-8 Sequence 8, Appli

#### ALIGNMENTS

##### RESULT 1

US-09-023-655-694  
; Sequence 694, Application US/09023655  
; Patent No. 6607879  
; GENERAL INFORMATION:  
; APPLICANT: Cocks, Benjamin G.  
; APPLICANT: Susan G. Stuart  
; APPLICANT: Jeffrey J. Seilhamer  
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE  
; TITLE OF INVENTION: EXPRESSION  
; NUMBER OF SEQUENCES: 1508  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
; STREET: 3174 PORTER DRIVE  
; CITY: PALO ALTO  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
; CURRENT APPLICATION NUMBER: US/09/023,655  
; FILING DATE: HEREWITH  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Zeller, Karen J.  
; REGISTRATION NUMBER: 37,071  
; REFERENCE/DOCKET NUMBER: PA-0001 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (650) 855-0555  
; TELEFAX: (650) 845-4166  
; INFORMATION FOR SEQ ID NO: 694:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 567 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: MMLR2DT01  
; CLONE: 476160  
US-09-023-655-694

Query Match 19.0%; Score 560.8; DB 4; Length 567;

Best Local Similarity 99.1%; Pred. No. 1.9e-109;  
Matches 562; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1668 CACCATCTTACAGGCTTCTCCTCGAGAGCCGCGCGGACACCCCTGAGCTCAA 1727  
Db 1 CACCATCTTACAGGCTTCTCCTCGAGAGCCGCGCGGACACCCCTGAGCTCAA 60

QY 1728 GCCACCGCTAGTGGCTTTTCAACATTCCTCCAGCCTTCCAGTGCAGTCCGCCAC 1787  
Db 61 GCCACCGCTAGTGGCTTTTCAACATTCCTCCAGCCTTCCAGTGCAGTCCGCCAC 120

QY 1788 TGACCTTCACTCACACGACAGATGAGGAGGCACTTGGAGTGTGGTGCC 1847  
Db 121 TGACCTTCACTNCAACAGGACAGATGAGGAGGCACTTGGAGTGTGGTGCC 180

QY 1848 CAGGACGGTCCCTCCAGCTCAACAGTGGGCTTGAAGTGGACAGGTTAATGTCTCCAGAGTGA 1907  
Db 181 CAGGACGGTCCCTCCAGCTCAACAGTGGGCTTGAAGTGGACAGGTTAATGTCTCCAGAGTGA 240

QY 1908 CACTGAGGAGCAGCATTTACAGGCTGAGTGTGTTTCCGGAGTCTGTCACAGGCC 1967  
Db 241 CACTGAGGAGCAGCATTTACAGGCTGAGTGTGTTTCCGGAGTCTGTCACAGGCC 300

QY 1968 ACACGCTCACTTGACTCATGCTGCTAAGATGACACCGCACACCCATACAACTACAA 2027  
Db 301 ACACGCTCACTTGACTCATGCTGCTAAGATGACACCGCACACCCATACAACTACAA 360

QY 2028 GGGCCACAAAGCACTGCTGGGTAGCTTTTCCACAGACATAATATAGTCCATCTGCAAT 2087  
Db 361 GGGCCACAAAGCACTGCTGGGTAGCTTTTCCACAGACATAATATAGTCCATCTGCAAT 420

QY 2088 CACAGCAGATGACGAGTAAACGACCACTCCCTGGATCTGACGCCACACAGTGGGAG 2147  
Db 421 CACAGCAGATGACGAGTAAACGACCACTCCCTGGATCTGACGCCACACAGTGGGAG 480

QY 2148 TCTGGCTGTCACCTTACAAAGCCACAGAAAGGCGCACATGTTACAGCTACACGCC 2207  
Db 481 TCTGGCTGTCACCTTACAAAGCCACAGAAAGGCGCACATGTTACAGCTACACGCC 540

QY 2208 TCTCCATTCATGAACTTCTCAGTGTC 2234  
Db 541 TCTCCATTCATGAACTTCTCAGTGTC 567

RESULT 2

US-08-277-031B-8  
; Sequence 8, Application US/08277031B  
; Patent No. 6620593  
; GENERAL INFORMATION:  
; APPLICANT: Hayashi, Koji  
; APPLICANT: Sakaki, Toshiyuki  
; APPLICANT: Yabusaki, Yoshiyasu  
; APPLICANT: Komai, Koichiro  
; APPLICANT: Kaneko, Hideo  
; APPLICANT: Nakatsuka, Iwao  
; TITLE OF INVENTION: METHOD FOR SAFETY EVALUATION OF  
; TITLE OF INVENTION: CHEMICAL COMPOUND USING RECOMBINANT YEAST EXPRESSING  
; NUMBER OF INVENTIONS: 42  
; NUMBER OF SEQUENCES: 42  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Birch, Stewart, Kolasch & Birch  
; STREET: P.O. Box 747  
; CITY: Falls Church  
; STATE: Virginia  
; COUNTRY: USA  
; ZIP: 22040-0747  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.5inch, 1.44MB  
; COMPUTER: IBM PC  
; OPERATING SYSTEM: Dos 5.0  
; SOFTWARE: Wordperfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/277,031B

FILING DATE: 19-JULY-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP-201120/1993  
; APPLICATION NUMBER: JP-180246/1993  
; APPLICATION NUMBER: JP-208279/1993  
FILING DATE: 20-07-1993  
FILING DATE: 21-07-1993  
FILING DATE: 30-07-1993  
ATTORNEY/AGENT INFORMATION:  
; NAME: Raymond C. Stewart  
; REGISTRATION NUMBER: 21,066  
; REFERENCE/DOCKET NUMBER: 20-3530P  
TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703-205-8000  
; TELEFAX: 703-205-8050  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1485  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
US-08-277-031B-8

Query Match 16.1%; Score 475.2; DB 4; Length 1485;  
Best Local Similarity 59.6%; Pred. No. 3.2e-91; Indels 6; Gaps 2;  
Matches 839; Conservative 0; Mismatches 563;

QY 378 GACCAGGCGCCGAGGCCACCTGCCCCCGGCGCCACCGCTACCACTCTGGGAAACCT 437  
Db 78 GAGGAGAGCAAGGGGAGCTGCTCCGCGGACCCACCCCATTTGCCCTTATTGAAACTA 137

QY 438 CTTGAGAGTACGCGCGGCGCTGTATTTCAGGGCTCATCGGCTAGTAAAGTACGG 497  
Db 138 CTTGAGAGTAAACACAGAGCAGATGTACAACTCCCTCATGAAGATCAGTGAAGCTATGG 197

QY 498 ACCGGTGTTCACCATCTACCTGGGACCTGGCGGCTGTGTGTCTCTGGGCGAGGA 557  
Db 198 CCGCGTGTTCACCATCTACCTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 254

QY 558 GGTGTGCGGAGGCGCCCTGGGAGGTGAGGTGAGGAGTTCAGCGCGCGGCGGAAACCGTAGC 617  
Db 255 TGCGGTGAGGAGGCTCTGTTGAGCCAGGCTGAGGAGTTCAGCGCGGCGGCGGCGGCGGCGG 314

QY 618 GATGCTGAGGAGGACTTTTGTATGGCCATGGGTTTCTTCTCCAAAGGCGGAGCGGTGGAG 677  
Db 315 CACCTTCGACTGGGTCTTCAAAGGCTATGGCGTGTGTTTTCAGCAACGCGGAGCGCGCAA 374

QY 678 GCAGCTGAGGAGGTTTACCATGCTTCTGCGGAGCTTGGGCTGCGGAGCGGAGGAGG 737  
Db 375 GCAGCTGCGGCGGTTTCTCATCGCACCTGCGGAGCTTGGGCTGCGGAGCGGAGGAGG 434

QY 738 CGAGGAGTGTATCCAGCGGAGGCGCGGCTGTCTGTGTGAGAGATTCAGGCGGAGCAGAGG 797  
Db 435 CGAGGAGGCGCATCCAGGAGGAGGCGGCGGCTTCTCATCGACGCGCTCCGCGGCGCACTGGCGG 494

QY 798 AGCGCCATTCGATCCCTCCCTGCTGTGCGCCAGGCCACCTCCAAAGTGTCTGCTCCCT 857  
Db 495 CGCCAAATATCGATCCCACTTCTTCTGAGCGCGCACAGTCTTCAATGTCTATGAGTCCAT 554

QY 858 CTTCTTTGGCTCCGCTTCTCTATGAGGATGAGGAGTTCAGGCGGCTGTCTCGGCGAGC 917  
Db 555 TGCTTTTGGGACCGCTTGTGACTATGAGGACAAAGAGTTCCTGTCTACTGTGCGCATGAT 614

QY 918 TGGTGTACCTGCTCGGAGTTCAGCTCCAGGCGGCGTTCAGACTAGAGATTTCTCTG 977  
Db 615 GCTAGGAATCTTCCAGTTCAGTCAACCTCCAGGCGGCGAGCTCTATGAGATGTTCTCTTC 674

QY 978 GTTCTCTGGCGGCGCTGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1037  
Db 675 GGTGATGAAACACTCTGCGAGGACACAGGAGGCGGCTTTCAGTGTCTGCAAGGCGCTGGA 734

1038	QY	TGCTTCACTCGGAGTGCAGCAGCACGAGGAACCTGATGCTTTGGGCCCCGC	1097
735	Db	GGACTTTTCATAGCCAAGAAGGTGGAGCAACAACAGCGCAGCTGGATCCCAATTCGC	791
1098	QY	ACGTGACCTTTGTGATGCCCTTCCTGCTGAAGATGGCACAGAGGAAACAACCCAGGCAC	1157
792	Db	ACGGGACTTCATTTGACTTCCTTTTCATCCGCATGCAGGAGGAGGAAGAACCCCAACAC	851
1158	QY	AGAATTCAACAACAAGAACATGCTGATGACAGATCATTTATTGTGTTTCTCTGGGACGAT	1217
852	Db	GGAGTTCTACTTGAAAAACCTGGTGATGACCAAGTTGAACCTCTTCATTTGGGGGACCGA	911
1218	QY	GACGGTCAGACACACGGTCGGCTATACCTTCCTGCTCCTGATGAATACCCCTCATGTCCA	1277
912	Db	GACCGTCAGACACACCCCTGGCGCTATGGCTTCTTGCTCTCATGAAGCACCCAGAGTTGA	971
1278	QY	AAAGTGGTACGTGAGGAGCTGAATCGGAGCTGGGGCTGGCCAGGCACCAAGCCTAGG	1337
972	Db	GGCCAAAGTCCATGAGAGATTCACAGAGTGTGCGCAAGACCGCAGGCCCAAGTTTGA	1031
1338	QY	GGACCGTACCCCGCTCCCTTTACACCGACGGGTTCCTGCATGAGCGCAGCGGCTGTGGC	1397
1032	Db	GGACCGGGCCAAGATGCCCTACATGGAGGACGTGATCCACAGATCCAAAGATTTGGAGA	1091
1398	QY	GCTGGTGCCCATGGGAATACCCGCAACCTTCATGCGGAACACCCGCTTCGAGGGTACAC	1457
1092	Db	CGTGATCCCCATAGATTTGGCCCGCAGAGTCAAAAAGGACACCAAGTTTCGGGATTTCT	1151
1458	QY	CTTGCCCCAGGGCAGGAGGTCTTCCCGCTCCTTGGCTCCATCCTGATGACCCCAACAT	1517
1152	Db	CTTCCCTAAGGGCACCGAAGTGTACCCCTATGCTGGGCTCTGTGCTGAGAGACCCCAAGTTT	1211
1518	QY	CTTCAAGCACCCAGAAAGATTCAACCCAGACCGTTTCTTGGATGACAGTGCAGTTCAG	1577
1212	Db	CTTCTCAACCCCGGACCTTCAATCCACAGACTTCTTGAATGAGAAGGGGAGTTTAA	1271
1578	QY	GAAGCATGAGGGTTCCTGCGCTTCTCTTAGGAAAGCGTGTGTGCTTTGGAGAGGCGCT	1637
1272	Db	GAAGAGTGTATGCTTTTGTGCGCTTTTCCATCGGAAAGCGAACTGTTCGGAGAAAGCGCT	1331
1638	QY	GGCAAAAGCGGAGCTCTTCTCTTCTTTCACCACCATCTTACAAGCCTTCTCCTTGGAGAG	1697
1332	Db	GGCCAGAATGGAGCTCTTCTCTTCTTTCACCACCGTCATGCAAGAACTTCGCGCTCAAGTC	1391
1698	QY	CCCGTCCGCGCGGACACCTGAGCCTCAAGCCACCGTCAGTGGGCTTTTCAAATCCC	1757
1392	Db	CTTCCAGTCACCTAAGGACATTCAGCTGTCCCCCAGACAGCTGGGCTTTCACAGATCCC	1451
1758	QY	CCGAGCTTCAGCTGCAAGTCGTTCC	1785
1452	Db	ACGAAACTACACCATGAGCTTCTGTCCC	1479

## RESULT 3

```

US-08-277-031B-9
; Sequence 9, Application US/08277031B
; Patent No. 6620593
; GENERAL INFORMATION:
; APPLICANT: Hayashi, Koji
; APPLICANT: Sakaki, Toshiyuki
; APPLICANT: Yabusaki, Yoshiyasu
; APPLICANT: Komai, Koichiro
; APPLICANT: Kaneko, Hideo
; APPLICANT: Nakatsuka, Iwao
; TITLE OF INVENTION: METHOD FOR SAFETY EVALUATION OF
; TITLE OF INVENTION: CHEMICAL COMPOUND USING RECOMBINANT YEAST EXPRESSING
; TITLE OF INVENTION: HUMAN CYTOCHROME P450
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch
; STREET: P.O. Box 747
; CITY: Falls Church
; STATE: Virginia

```











Db 92 GGAAGTGCCTCGGACCCACCCATTGCCCTTCAATGGAACTACCTCCAGCTGAACA 151  
QY 452 CGGGGCGCTGTATTCAGGGCTCATGCGCTCAGTAAGAAAGTACGGACCGGTTCACCA 511  
Db 152 CAGAGCAGATGTACAACTCCCTCATGAAGATCAGTGAGCGCTATGGCCCTGTGTTCACCA 211  
QY 512 TCTACCTGGGACCTCGGCGCTGTGGTGTCTGTGGGAGGAGGCTGTGCGGAGG 571  
Db 212 TTCACTTGGGGCCCGCGCG--GTGCTGTGTGTGCGGACATGATGCCGTCAAGGAGG 268  
QY 572 CCTTGGAGCTCAGGCTGAGGAGTTTCAGCGCGCGGGAACCGTAGCGATGCTTGAAGGGA 631  
Db 269 CTCGTGTGACCAAGCTTGAGAGTTTCAGCGGCGAGCGAGCAGGCGCACTTCGATGCG 328  
QY 632 CTTTGTATGCCATGGGTTTTCTTCCAAAGGGGAGCGGTGGAGGAGCTGAGAAAT 691  
Db 329 TCTTCAAAGGCTATGCGGTGGCGTTTCAGCAACCGGGAGCGCGCAAGCAGCTCCGCGCT 388  
QY 692 TTACCATGCTTGTCTCGCGGACCTGGGATGGGAGCGAGAGCGAGGAGCTGATCC 751  
Db 389 TCTCCATCGCCACCTTAAGGGTTTTTGGCGTGGGCAAGCGCGCATCGAGGAACGCATCC 448  
QY 752 AGCGGAGCGCCGCTGTCTGTGGAGACATTCAGGGGACAGAAAGCGCCATTCGATC 811  
Db 449 AGGAGAGGGCGGCTTCTCTCATCGAGCCCTCGGGGCGAGCAGCGGCGCAATATCGATC 508  
QY 812 CTTCTCTGTGTGGGCGGCGGCGACCTCCAACTAGTCTGTCTCTCTCTCTTTGGGCTCC 871  
Db 509 CCACCTTCTCTGTAGCGGCGACAGTCTCCAACTCATCAGTCTCAATGTCTTTGGGAGC 568  
QY 872 GCTTCTCTATGAGGATAGGAGTTCCAGCGCTGTCTCGGCGAGCTGTGGTACCTGC 931  
Db 569 GCTTTGACTATGAGCAAGAGTTCTGTCTGCTGTGGCGATGATGCTGGGAAGGTTCC 628  
QY 932 TGGGAGTACCTCCAGGGGGTTCAGACCTTACAGATGTTCTCTGCTTCTCGGCGCC 991  
Db 629 AGTTCAGGGAACCTCCAGGGGCGAGTCTATGAGATGTTCTCTCGGTGATGAACACC 688  
QY 992 TGGCAGGCGCCCAAGAGAGCTCTCCACAGCTCAGACCTTGGCTGCGCTTACAGTCC 1051  
Db 689 TGGCAGGACACAGCAA-CAGGCGCTTAAAGAGCTGCAAGGGCTGGAGGACTTCATCGCCA 747  
QY 1052 GGCACTGTCAGACGACAGGGGAACTGATGCTTTCGGGCGCGGCGAGCTGTGCTG 1111  
Db 748 AGAAGGTGAGACAAACAGCGGACGCTGA---TCCCAATCCCGGCGGATTCATCG 804  
QY 1112 ATGCGCTTCTGTGAGAGTGGCACAGGAGAAACAAACCCAGGCACAGAAATTCACCAACA 1171  
Db 805 ACTCTTCTCATCCGATCGAGAGGAGGAGAAACCCCAACACAGAGTTCATTTGA 864  
QY 1172 AGAATATGCTGATGACAGTCAATTAATTTGCTGTGGAGATGACGCTGAGCACCA 1231  
Db 865 AGAACCTGTGTATGACACCCCTGAACTCTTCTTGGGGCACTGAGACCGCTGAGCACCA 924  
QY 1232 CGCTCGGCTATACCTCTCTCTCTGATGAATACCTCATGTCCTCAAAAGTGCTGCTG 1291  
Db 925 CCTCGGCTACGGTTTCTCTGCTCTATGAAGACCCAGAGGTGGAGGCGCAAGTTCATG 984  
QY 1292 AGGAGCTGATCGGGAGCTGGGGCTGGCGAGCACCAAGCCTTAGGGGACCGTACCGGCC 1351  
Db 985 AGGAGATTGACAGATGATCGGCAAGAACCGGAGCCCAAGTTTGAAGGACCGGCGCAAGA 1044  
QY 1352 TCCCTTACAGGAGGCTGTGATGAGGCGGAGCGGCTGCTGGCGCTGTGCGCCATGG 1411  
Db 1045 TGGCCTACAGAGGCGAGTATCCACAGATCCAAAGATTGGAGACATGCTCCCCATGG 1104  
QY 1412 GAATACCCCGACCTCATCGGACACCGCTTCCGAGGCTACACCTGCGCCAGGCA 1471  
Db 1105 GTTTGGGCGCACAGGTTCAACAGGACACCAAGTTTCGGGATTTCTCTCTCTTAAGGGA 1164  
QY 1472 CGGAGGTCTTCCCGCTCTTGGCTCCATCTGCAATGACCCCAACATCTTCAAGCAACCCAG 1531

Db 1165 CTGAAGTGTTCCTATGCTGGGCTCCGAGCTGAGAGACCCAGGTTCTTCTCCAAACCCC 1224  
QY 1532 AAGAGTTCAACACAGACGCTTTCCTGGATGAGATGAGAGCGTTTCAGGAAGCATGAGCGT 1591  
Db 1225 AGGACTGCACTCCCGAGCACTTCTCTGGATGAGAGGGCAGTTTAAAGAGTATGCTT 1284  
QY 1592 TCTGCGCTTCTCTTAGGAAAGCGTGTCTGCTTGGAGAGGCGCTTGGCAAAAGCGGAGC 1651  
Db 1285 TTGTCGCTTTCATCGAAAGCGTACTGTGTTGGAGAGGCGCTGGCCAGAAATGAGC 1344  
QY 1652 TCTTCTCTTCTTCAACCATCTTCAAGCTTCTTCCCTGGAGAGCGGTCGCCCGCG 1711  
Db 1345 TCTTCTCTTCTTCAACCATCTTCAAGCTTCTTCCCTTCAAGTCCCTCGCTCGCCTA 1404  
QY 1712 ACACCTGAGCTCAAGCCACCGTCAGTGGCTTTTCAACATTCGCCAGGCTTCCAGC 1771  
Db 1405 AGATATCAGCTGTCCCCCAAACACAGTGGGCTTTCAGCATCCCAAGAACTACACA 1464  
QY 1772 TGCAAGTTCGCTCCC 1785  
Db 1465 TGAGCTTCTGCTCCC 1478

RESULT 8

US-09-023-655-1061  
; Sequence 1061, Application US/09023655  
; Patent No. 6607879  
; GENERAL INFORMATION:  
; APPLICANT: Cocks, Benjamin G.  
; APPLICANT: Susan G. Stuart  
; APPLICANT: Jeffrey J. Seilhamer  
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE  
; TITLE OF INVENTION: EXPRESSION  
; NUMBER OF SEQUENCES: 1508  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: INCYTE PHARMACEUTICALS, INC.  
; STREET: 3174 PORTER DRIVE  
; CITY: PALO ALTO  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM: disk  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
; CURRENT APPLICATION DATA:  
; FILING DATE: HEREWITH  
; APPLICATION NUMBER: US/09/023,655  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Zeller, Karen J.  
; REGISTRATION NUMBER: 37,071  
; REFERENCE/DOCKET NUMBER: PA-0001 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (650) 855-0555  
; TELEFAX: (650) 845-4166  
; INFORMATION FOR SEQ ID NO: 1061:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1825 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: GENBANK  
; CLONE: g181357  
; US-09-023-655-1061

Query Match 14.6%; Score 429.4; DB 4; Length 1825;



CLONE: 1749882  
US-09-023-655-459

Query Match 11.1%; Score 328; DB 4; Length 341;  
Best Local Similarity 99.1%; Pred. No. 2.3e-60;  
Matches 339; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1075 AACCTGGATGCTTGGGGCCCCCAGTGACCTTGTGATGCTTCTCTGTAAGATGCA 1134  
Db 1 AACCTGGATGCTTGGGGCCCCCAGTGACCTTGTGATGCTTCTCTGTAAGATGCA 60

QY 1135 CAGGAGGAACAAACCCAGGACAGAAATTCACCAACAGAAATGCTGATGACAGTCATT 1194  
Db 61 CAGGAGGAACAAACCCAGGACAGAAATTCACCAACAGAAATGCTGATGACAGTCATT 120

QY 1195 TATTGCTGTTTTGCTGGGACGATGACGGTCAGCACACCGTTCGGCTATACCTTCCTGCTC 1254  
Db 121 TATTGCTGTTTTGCTGGGACGATGACGGTCAGCACACCGTTCGGCTATACCTTCCTGCTC 180

QY 1255 CTGATGAATACCTCATGCTCCAAAGTGGGTACGTGAGGAGCTGAATCGGAGCTGGGG 1314  
Db 181 CTGATGAATACCTCATGCTCCAAAGTGGGTACGTGAGGAGCTGAATCGGAGCTGGGG 240

QY 1315 GCTGGCCAGCACCAAGCCTAGGGGACCGTACCGGCTCCCTTACACCGACGCGTTCTG 1374  
Db 241 GCTGGCCAGCACCAAGCCTAGGGGACCGTACCGGCTCCCTTACACCGACGCGTTCTG 300

QY 1375 CATGAGGCGACGCGGCTGCTGGCGCTGGTGCCCATGGGAATA 1416  
Db 301 CNTGAGGCGCA-CGGCTGCTGGCGCTGGTGCCCATGGGAATA 341

RESULT 10  
US-08-277-031B-15  
; Sequence 15, Application US/08277031B  
; Patent No. 6620593  
; GENERAL INFORMATION:  
; APPLICANT: Hayashi, Koji  
; APPLICANT: Sakaki, Toshiyuki  
; APPLICANT: Yabusaki, Yoshiyasu  
; APPLICANT: Komai, Koichiro  
; APPLICANT: Kaneko, Hideo  
; APPLICANT: Nakatsuka, Iwao  
; TITLE OF INVENTION: METHOD FOR SAFETY EVALUATION OF  
; TITLE OF INVENTION: CHEMICAL COMPOUND USING RECOMBINANT YEAST EXPRESSING  
; TITLE OF INVENTION: HUMAN CYTOCHROME P450  
; NUMBER OF SEQUENCES: 42  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Birch, Stewart, Kolasch & Birch  
; STREET: P.O. Box 747  
; CITY: Falls Church  
; STATE: Virginia  
; COUNTRY: USA  
; ZIP: 22040-0747  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.5inch, 1.44MB  
; COMPUTER: IBM PC  
; OPERATING SYSTEM: Dos 5.0  
; SOFTWARE: Wordperfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/277,031B  
; FILING DATE: 19-JULY-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP-201120/1993  
; APPLICATION NUMBER: JP-180246/1993  
; APPLICATION NUMBER: JP-208279/1993  
; FILING DATE: 20-07-1993  
; FILING DATE: 21-07-1993  
; FILING DATE: 30-07-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Raymond C. Stewart  
; REGISTRATION NUMBER: 21,066

REFERENCE/DOCKET NUMBER: 20-3530P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-205-8000  
TELEFAX: 703-205-8050  
TELEX:  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1473  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
US-08-277-031B-15

Query Match 11.0%; Score 324; DB 4; Length 1473;  
Best Local Similarity 53.0%; Pred. No. 2.8e-59;  
Matches 740; Conservative 0; Mismatches 650; Indels 6; Gaps 2;

QY 389 GAGGCCACCTGCTCCCCCGGGCCACCGCGCTACCTGCTGAGTAAGTACGACCGCTGTTCA 508  
Db 77 GAGGAAAACTCCCTCTCGGCCCACTCTCTCCAGTGATTGGAAATATCTACAGATAG 136

QY 449 GGGCCGGGGCGCTGTATTTCAGGGCTCATCGGCTGATGTAAGTACGACCGCTGTTCA 508  
Db 137 ATATTAAAGGATGTCAGCAAAATCCTTAACCAATCTCTCAAAAATCTATGGCCCTGTGTTCA 196

QY 509 CCATCTACTCTGGACCCCTGGCGGCTGTGGTGGTCTCTGGTGGGAGGAGGCTGTGCGGG 568  
Db 197 CTCTGTATTTTGGCCCTCGAGCCCA---TGGTGGTGTGTCATGGATATGAAGTGTGAAG 253

QY 569 AGGCCCTGGAGGTTCAGGCTGAGGAGTTCAGCGCCCGGGGAAACCGTAGCGATGCTGGAAG 628  
Db 254 AAGCCCTGATTGATCTTGGAGAGGAGTCTTCTGGAAGAGGCCATTTCCCACTGGCTGAAA 313

QY 629 GGACTTTTGTATGGCCATGGGTTTCTCTCCAAAGGGGAGCGGTGAGGAGCTGAGGA 688  
Db 314 GAGCTAACAGAGGATTTGGAAATCGTTTTTTCAGCAATGGAAGAGATGGAAGGAGATCGGC 373

QY 689 AGTTTACCATGCTTGTCTGCGGGAACCTGGGATGGGGAAGCGAGAGGAGGAGCTGA 748  
Db 374 GTTCTCCCTCATGAGCTGCGGAATTTTGGGATGGGGAAGAGGAGCATTTAGGACCGTG 433

QY 749 TCAGGCGAGGCGCGGTGTCTGTTGGAGACATTCAGGGGGAAGAGAGACGCCCATTCG 808  
Db 434 TTCAAGAGGAAGCGCGCTGCTTGTGGAGAGTTGAGAAAAACCAAGGCTTCAACCTGTG 493

QY 809 ATCCCTCCCTGCTGCTGCGCCAGGCCACCTCCAACTAGTCTGCTCCCTCTCTTTGGCC 868  
Db 494 ATCCCACTTTCATCCTGGGCTGTGCTCCCTGCAATGTGATCTGCTCATTTATTTCCAGA 553

QY 869 TCCGCTTCTCTATGAGGATAAGGAGTTCCAGGCGGTGTCGCGGCGAGCTGGTGGTACCC 928  
Db 554 AAGGTTTCGATTATAAAGATCAGCAATTTCTTAACTTGATGGAATAATTTGAATGAAAA 613

QY 929 TGCTGGAGTTCAGCTCCAGGGGGGTTCAGACCTACAGATGTTCTCTGTTCTCTGGGC 988  
Db 614 TCAGGATTTAGACACCCCTCGGATCCAGATATGCAATAATTTTCCACATATCATTGAT 673

QY 989 CCCTGCCAGGCCCCCAAGACAGCTCTCCACCACTGTCAGCACCTTGGCTGCTTCACAG 1048  
Db 674 ATTTCCGGGAACCCATAACAAATTACTTAACAACTTGTCTTTTATGGAAGTATATTT 733

QY 1049 TCCGGCAGGTGTCAGCAGCACCGGGGAACCTGGATGCTTCCGGCCCCCGGACGTGACCTTG 1108  
Db 734 TGGAGAAAGTAAAGAAACACCAAGAATCGATGGA---CATCAACAAACCTTCGGGACTTTA 790

QY 1109 TCAGTGCCTTCTGCTGAGATGGCAGAGGAAACAAACCCAGGACAGAGATTCACCA 1168  
Db 791 TTGATTGCTTCTGATCAAAATGGAGAGAAAGCAACCAACAGTCTGAAATCTACTA 850

QY 1169 ACAAGAACATGCTGATGACAGTCAATTTATTTGCTGTTTGTCTGGGACGATACCGTTCAGCA 1228  
Db 851 TTGAAAACTTGGTAATCACTGACGCTCACTTACTTGGAGCTGGGACAGAGACAAAGCA 910







```
QY 1049 TCGGAGGTGTCAGCAGCAGCAGGGAACCTGGATGCTTCGGGCCCGCAGCTGACCTTG 1108
Db 746 TGGAAAGTAAAGAAACACAGAAATCAATGGA---CATGAACACCTCAGGACTTTA 802
QY 1109 TCGATGCTTCTCTGCTGAAGATGGCAGGAGGAACAAACCCAGGACACAGAAATTCACCA 1168
Db 803 TTGATGCTTCTCTGATGAAATGGAAGGAAGACACACCAACCATCTGATTTACTA 862
QY 1169 ACAGAACATGCTGATGACAGTCAATTTATTTGCTGTTTGTGGACATGACGTCAGCA 1228
Db 863 TTGAAAGCTTTGAAACACACTGCGATTTGACCTGTTTGGAGCTGGGACAGACACAGCA 922
QY 1229 CCACGCTGCGCTATACCTCTGCTGCTGATGAATACCTCTCATGTCCAAAGTGGTAC 1288
Db 923 CAACCTTGAGATATGCTCTCTCTGCTGAAGACCCAGAGGTACAGCTAAAGTCC 982
QY 1289 GTGAGGAGCTGAATTCGGAGCTGGGGGCTGGCCAGGACCAAGCCTTAGGGGACCGTACCC 1348
Db 983 AGCAAGAGATTGAACGTGTGATTGGCAGAAACCGGAGCCCTGCATGCAAGACAGGACC 1042
QY 1349 GCCTCCCTTACACGAGCGGGTCTGATGAGGGGACAGCGGCTGCTGGCGCTGGTCCCA 1408
Db 1043 ACATGCCCTACACAGATGCTGTGTGACGAGGTCCAGAGATACCTTTGACCTTCTCCCA 1102
QY 1409 TGGGAATACCCGACCCCTCATGCGGACCAACCGCTTCCGAGGTACACCTGCCCGAG 1468
Db 1103 CAGCTGCCCCATGAGTGAATTTCAAGAACTATCTCATTTCCCAAGG 1162
QY 1469 GCAGGAGGTCTTCCCTCTGCTGCTGATGATGATGATGATGATGATGATGATGATGAT 1528
Db 1163 GCACCAACCATTAATTTCCCTGACTCTCTGCTACATGACACAAAGAAATTTCCCAAC 1222
QY 1529 CAGAGAGTTCAACCAAGCGTTTCTGATGAGGAGTACAGTACAGTACAGTACAGTACAG 1588
Db 1223 CAGAGATGTTGACCTCTACATCTTCTGATGAAGGTGCAATTTTAAAGAAAGTAAAT 1282
QY 1589 CGTTCTGCGCTTCTCTTAGGAAGCGTCTGCTGCTGCTGAGAGGCGCTGCAAGAGCG 1648
Db 1283 ACTTATGCTTCTCAGCAGGAAGCGATTTGTTGGGAGAGCGCTGCGCGGATGG 1342
QY 1649 AGCTTCTCTCTTCTTCAACCACTCTCAAGCCTTCTCCCTGGAGAGCGCTGCGCGC 1708
Db 1343 AGCTGTTTTTATCTGACCTCCATTTTACAGAACTTTAACTGAAATCTCTGTTGACC 1402
QY 1709 CGGACACCTGACCTCAAGCCACCGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1768
Db 1403 CAAAGAACCTTGACACCACTCCAGTTGTCAATGGTTTGTGCTGCTGCTGCTGCTGCT 1462
QY 1769 AGCTGCAAGTCCCTCC 1784
Db 1463 AGCTGCTTCTATCC 1478
```

```
RESULT 13
US-09-023-655-1056
; Sequence 1056, Application US/09023655
; Patent No. 6607879
;
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
; TITLE OF INVENTION: EXPRESSION
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
```

```
;
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/023,655
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0001 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1056:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1854 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: g181301
; US-09-023-655-1056

Query Match 11.0%; Score 322.4; DB 4; Length 1854;
Best Local Similarity 52.9%; Pred. No. 6.7e-59;
Matches 739; Conservative 0; Mismatches 651; Indels 6; Gaps 2;

QY 389 GAGGCCACCTGCCCCCGGCCCGCCGCTACCTGCTGGGAAACCTCTCTGACGCTAC 448
Db 89 GAGGAAACCTCCCTCTCTGCGCCCACTCTCTCCAGTATTGGAAATATCTACAGATAG 148
QY 449 GCGCCGGGGCGCTGTATTTAGGGCTCATGCGGCTAGTAAAGTACGAGACCGGTGTCA 508
Db 149 GTATTAGGACATCAGCAATCTTAAACCAATCTCTAAAGGTCTATGGCCCTGTGTCA 208
QY 509 CCATCTACCTGGACCTTGGCGGCTGTGGTCTCTGCTGGGAGGAGCTGTGGGG 568
Db 209 CTCTGATTTTGGCCT---GAAACCCATAGTGTGTGCTGATGATGATGATGATGATG 265
QY 569 AGCCCTGAGGAGTCAAGGTGAGGAGTTCAGCGCGCGGGAACCGTAGCGATGCTGAAG 628
Db 266 AAGCCCTGATTGATCTTGGAGAGAGTCTTCTGGAAGAGGCAATTTCCCACTGGTGAAA 325
QY 629 GGAATTTTGTATGGCCATGGGGTTTTCTCTCCAAACCGGGAGCGGTGAGGAGCTGAGGA 688
Db 326 GAGCTAACAGAGGATTTGGAATTTGTTTTCAGCAATGGAAGAAATGGAAGAGATCCGCG 385
QY 689 AGTTTACCATGCTGCTCTGCGGACCTGGGATGGGAGCGAGAGGAGGAGCTGA 748
Db 386 GTTCTCTCTCTATGACGCTGCGGAATTTTGGGATGGGAGAGAGGAGCATTTGAGGACCGTG 445
QY 749 TCCAGGCGGAGGCGCGGTGTCTGTTGGAGACATTTCCAGGGGACAGAGGACGCCCATTCG 808
Db 446 TTCAAGAGGAGAGCGCGCTGCTCTGTTGGAGAGTGTGAGAAAAACCAAGGCTCACCCTGTG 505
QY 809 ATCCCTCCCTGCTGCTGGGCGCAGGCGCACTCCAAACGTAGTCTGCTCCCTCTCTTTGGCC 868
Db 506 ATCCCACTTTTCATCTCTGGGCTGTGCTCCCTGCAATGTGATCTGCTCCATTAATTTCCATA 565
QY 869 TCGGCTTCTCTATGAGGATAGGAGTTCAGGCGGCTGCTCCGGGAGCTGTTGGTACCC 928
Db 566 AAGTTTGTATTATAAGATCAGCAATTTCTTAACCTTAATGGAAGATTTGAATGAAGAAACA 625
QY 929 TGCTGGGAGTCACTCCAGGGGGGTGAGACCTTACGAGATGTTCTCTGCTGCTCTGCGGC 988
Db 626 TCAAGATTTTGAGAGCGCCCTGATCCAGATCTGCAATTAATTTTCTCTCTATCATTTGATT 695
```

```
QY 989 CCCTGCCAGGCCCCACAAGAGCTCTCCACCAAGTCAGCACCTTGGCTGCTTACAG 1048
D 986 ACTTCCCGGAACTCACAACAATTACTTAAAGCTTGTCTTTTAAAGTATATTT 745
QY 1049 TCCGCGAGGTGACAGCAGCAGGAGGAACTTGGATGCTTGGGCCCCGACGCTGACCTTG 1108
D 746 TGGAAAAAGTAAAGAACACCAAGAATCAATGGA--CATGAACAACCTCAGGACTTGA 802
QY 1109 TCGATGCTTCTGCTGAGATGGACAGGAGGAAACAAACCCAGGACAGAAATTCACCA 1168
D 803 TTGATGCTTCTGATGAATGGAGAGGAAAGACAAACCAACCATCTGAAATTTACTA 862
QY 1169 ACAAGAACATGCTGATGACAGTCAITTTATTTCTGTTTCTGGGACGATGACGGTCAGCA 1228
D 863 TTGAAGCTTGGAAAAACATGACGTGACTTGTGAGCTGGGACAGAGACCAAGCA 922
QY 1229 CACGGTGGGTATACCTCTCTCTGCTGATGAATAACCTCATCTGTCAAAGAGTGGGTAC 1288
D 923 CAACCTGAGATATGCTCTCTCTCTCTCTGCTGAGGACCCAGAGGTACAGCTAAAGTCC 982
QY 1289 GTGAGGAGCTGAATCGGAGCTGGGGCTGGCGGAGCCACCAAGCCTAGGGACCGTACC 1348
D 983 AGGAAGAGATTGAACGTGTGATGGCAGAAACCGAGGCCCTGCATGCAAGACAGAGCC 1042
QY 1349 GCCTCCCTTACACCAAGCGGTTCTGCAATGAGCGCAGCGGCTGCTGGCGCTGGTGCCA 1408
D 1043 ACATGCCCTACACAGATGCTGTGTGACAGAGTCCAGAGATACCTTGACCTTCTCCCA 1102
QY 1409 TGGGAATACCCGCAACCTCTATGCGGAGCACCCGCTTCCGAGGTACACCTGCCCCAGG 1468
D 1103 CCAGCTGCCCATGACGAGTACCTGTGACATTAAATTCAGAACTATCTCATTTCCCAAG 1162
QY 1469 GCAGGAGGTCTTCCCTCTCTGCTGCTCATCTGATGACCCCAACATCTTCAAGCACC 1528
D 1163 GCACAAACATATAATTTCCCTGACTTCTGTGTACATGACAAACAAAGAAATTTCCCAAC 1222
QY 1529 CAGAAAGTTCAACCAAGCAGCTTCTCTGATGACAGTGGAGCGTTTCAGGAAGCATGAGG 1588
D 1223 CAGAGATGTTGACCTCATCACTTCTGATGAGAGTGGCAATTTTAAAGAAAGTAAT 1282
QY 1589 CGTTCCTGCTCTCTCTAGGAAAGCTGTCTGCTTGGAGAGGCGCTGGCAAAAGCGG 1648
D 1283 ACTTCATGCTTCTCAGCAGGAAACGAAATTTGTGTGGGAGAGCGCTGGCGGCGATGG 1342
QY 1649 AGCTCTCTCTCTTCAACCACTCTACAGCTTCTCTGCTGAGAGCGCTGCCCCG 1708
D 1343 AGCTGTTTTTATCTCTGACCTCCATTTTACAGAACTTTAACTGAAATCTCTGTTGACC 1402
QY 1709 CGGACACCTTGAGCCTCAAGCCACCGCTGAGTGGGCTTTTCAACATTTCCCCAGCCTTCC 1768
D 1403 CAAAGAACCTTGACACCACTCCAGTTGTCATGTTTGGCTTCTGTGGCGCTTCTTACC 1462
QY 1769 AGCTGGAAGTGGTCC 1784
D 1463 AGCTGTGCTTCAATTC 1478
```

## RESULT 14

PCT-US95-05744-4

```
; Sequence 4, Application PC/TUS9505744
; GENERAL INFORMATION:
; APPLICANT: GOLDSTEIN, Joyce A.
; APPLICANT: ROWKES-SPARKS, Marjorie
; APPLICANT: DE MORALS, Sonia M.F.
; TITLE OF INVENTION: CLONING, EXPRESSION AND DIAGNOSIS OF HUMAN
; TITLE OF INVENTION: CYTOCHROME P450 2C19: THE PRINCIPAL DETERMINANT
; TITLE OF INVENTION: OF S-MEPHNYTOIN METABOLISM
; NUMBER OF SEQUENCES: 61
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: 379 Lytton Avenue
; CITY: Palo Alto
; STATE: California
```

```
; COUNTRY: US
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/05744
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/238,821
; FILING DATE: 06-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/201,118
; FILING DATE: 22-FEB-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/864,962
; FILING DATE: 09-APR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Dow, Karen B.
; REGISTRATION NUMBER: 29,684
; REFERENCE/DOCKET NUMBER: 15280-192-1-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 326-2400
; TELEFAX: (415) 326-2422
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1854 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; PCT-US95-05744-4
```

Query Match 11.0%; Score 322.4; DB 5; Length 1854;

Best Local Similarity 52.9%; Pred. No. 6.7e-59;

Matches 739; Conservative 0; Mismatches 651; Indels 6; Gaps 2;

```
QY 389 GAGGCCACCTGCCCCCGGCCCCGCGCTACCACTGCTGGGAAACCTCTCGAGCTAC 448
D 89 GAGGAAAACTCCCTCCTGGCCCCACTCCTCTCCAGTGTGTTGAAATATCTACAGATAG 148
QY 449 GSCCGCGGCGCTGTATTTCAGGGCTCATGCGGCTCAGTAAGAAGTACGACCGGTGTCCA 508
D 149 GTATTAAAGACATCAGCAAAATCCTTAACCAATCTCTCAAGAGTCTATGGCCCTGTGTTCA 208
QY 509 CCATCTACTCTGGACCCCTGGCGGCTGTGTGTGCTCTGTGTTGGGAGGAGGCTGTGCGG 568
D 209 CTCGTATTGTCCT--GAAACCCATAGTGTGCTCATGATATGAAGCAGTGAAG 265
QY 569 AGGCCCTGGAGGTGAGGTGAGGAGTTCAGCGCGCGGGGAAACCGTAGGATGCTGGAAG 628
D 266 AAGCCCTGATTGATCTTGAGAGGAGTTCCTGGAAGAGGCAATTTTCCCACTGGCTGAAA 325
QY 629 GGACTTTTGTATGGCCATGGGTTTTCTTCTCCAAACGGGAGCGGTGGAGGAGCTGAGA 688
D 326 GAGCTAACAGAGGATTTGGAATTTTTCAGCAATGGAAGAAATGGAAGAGATCCGCGC 385
QY 689 AGTTTACCATGCTTCTCTCGGGGACCTGGGCATGGGGAAGCGAAGAGCGAGGAGCTGA 748
D 386 GTTCTCCTCATGACGCTGCGGAATTTTGGCATGGGAAGAGGAGCATTTGAGGACCGTG 445
QY 749 TCCAGCGGAGCGCGGTGTCTGTGGGACATTTCCAGGGGACAGAAAGACGCCCATTCG 808
D 446 TTCAAGAGGAGCGCGCTGCTGTGGAGGAGTTCAGAAAAACCAAGGCTTCAACCTGTG 505
QY 809 ATCCCTCCTCTGCTGTGCGCCACAGGCACCTCCAAAGTAGTCTGTCTCTCTCTTTGGCC 868
D 506 ATCCCACTTTTCTCTGCGGCTGTGCTCCTCGCAATGTGATCTGTCTCTTATTTTCCATA 565
QY 869 TCCGCTTCTCTATGAGGATTAAGAGTTCACGCGCGTGGTCCGCGGAGCTGTGTGTACCC 928
```

```
Db 566 AACGTTTGTATTAAGATCAGCAATTTCTTAACCTTAATGGAAAGTTGAATGAAGAACA 625
QY 929 TGTGGAGTCAAGTCCCGAGGGGGTTCAGACCTACAGAGATGTTCTCTGTTCTCTGCGGC 988
Db 626 TCAAGATTTTGAGCAGCCCTGGATCAGATCTGCAATAATTTTCTCTATCATTTGATT 685
QY 989 CCCTGCCAGCCCCCAGACAGCTCTCTCCACAGCTCAGACCTTGGCTGCTTCACAG 1048
Db 686 ACTTCCCGGAATCACAACAATTAATCTTAAACGTTGCTTTATGAAAGTTATATTT 745
QY 1049 TCCGGAGTGCAGCAGCACCAGGGAACTTGGATGCTTCCGGCCCGCAGCTGACCTTG 1108
Db 746 TGGAAAGTAAAGAACACCAAGATCAATGGA--CATGAACAACCTCAGACTTGA 802
QY 1109 TCGATGCTTCTGCTGAAGATGCGCAGAGAGAAACAAACCCAGGACAGAAATTCACCA 1168
Db 803 TTGATGCTTCTGATGAAATGGAAGGAAAGCACAACCAACCATCTGAAATTTACTA 862
QY 1169 ACAAGACATGCTGATCAGACATCAATTTATTTGCTGTTGCTGGAGCATGACGTCAGCA 1228
Db 863 TTGAAGCTTGGAAACACTGCAAGTTGACTTGTGAGCTGGACAGACAGACAGCA 922
QY 1229 CACAGGTGCGGTATACCTCTCTGCTGATGAAATACCTCTCATGTCCAAAGTGGGTAC 1288
Db 923 CAACCTTGAGATATGCTCTCTCTCTGCTGAGCACCAGAGGTACAGCTTAAGTCC 982
QY 1289 GTGAGAGCTGAATCGGAGCTGGGGCTGGCAGGACCAAGCCTTAGGGAGCGTACCC 1348
Db 983 AGGAAGAGATTGAACGTGTGATGGCAGAACCGGAGCCCTGCTGATGCAAGACAGAGCC 1042
QY 1349 GCCTCCTTACCGAGCGGTTCTGATGAGCGGAGCGGCTGCTGGCCTGTTGCCCA 1408
Db 1043 ACATGCCCTACAGAGTGTGGTGACAGGTCCAGAGATACCTTGACCTTCTCCCA 1102
QY 1409 TGGGAATACCCCGACCTCATGCGGACCAACCGCTTCCGAGGTACACCTGCCCCAGG 1468
Db 1103 CCAGCCTGCCCATGAGTACCTGTGACATTAATTCAGAACTATCTCAITCCCAAG 1162
QY 1469 GCAGGAGTCTTCCCGCTCTGCTGCTCAATCTGATGACCCCAACATCTTCAAGCACC 1528
Db 1163 GCACAACCATATTAATTTCCGCTGACTTCTGCTGATGACATGACAAACAAAGAAATTTCCCAACC 1222
QY 1529 CAGAGAGTTCACCCAGCAGCTTCTGATGAGCGGAGTTCAGACGTTTCAAGAGCATGAGG 1588
Db 1223 CAGAGATGTTTGACCTCATCTTCTGATGAGGTGGCAATTTTAAAGAAAGTAAAT 1282
QY 1589 CGTTCTGCTCTCTCTTAGGGAAGCGTCTCTGCTTGGAGAGGGCTGGCAAAAGCGG 1648
Db 1283 ACTTCATGCTTCTCAGCAGGAAACGGATTTGTGAGGAGAGCCCTGGCCGATGG 1342
QY 1649 AGCTTCTCTCTTCAACCAACATCTCAAGCCTTCTCCCTGGAGAGCCCGTGGCCCG 1708
Db 1343 AGCTGTTTTTATTCCTGACCTCCATTTTACAGAACTTTAACTCTGATGAAATCTCTGTTGACC 1402
QY 1709 CGCAGACCTTGACCTCAAGCCACCGTCTGAGTGGCCTTTCAACATTTCCCGAGCTTCC 1768
Db 1403 CAAAGAACCTTGACACACTCCAGTTGTCATGTTTGGCTCTGTGCGCGCCTTCTTACC 1462
QY 1769 AGCTGCAAGTCCGTC 1784
Db 1463 AGCTGTCTCATTC 1478
```

## RESULT 15

```
US-08-194-981E-4
; Sequence 4, Application US/08194981E
; Patent No. 5886157
; GENERAL INFORMATION:
; APPLICANT: GUENGERICH, F. Peter
; APPLICANT: GUO, Zuyu
; APPLICANT: SANDHU, Punam
; APPLICANT: GILLAM, Elizabeth M. J.
```

```
; TITLE OF INVENTION: EXPRESSION AND PURIFICATION OF
; TITLE OF INVENTION: HUMAN
; NUMBER OF INVENTION: CYTOCHROME P450
; NUMBER OF SEQUENCES: 68
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NEEDLE & ROSENBERG, P.C.
; STREET: Suite 1200, 127 Peachtree Street, NE
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30303-1811
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/194,981E
; FILING DATE: February 10, 1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Elizabeth Selby
; REGISTRATION NUMBER: 38,298
; REFERENCE/DOCKET NUMBER: 22000.0022
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 688-0770
; TELEFAX: (404) 688-9880
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1419 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-08-194-981E-4
```

Query Match 10.9%; Score 321.8; DB 2; Length 1419;

Best Local Similarity 52.9%; Pred. No. 8.2e-59;

Matches 739; Conservative 0; Mismatches 652; Indels 6; Gaps 2;

```
QY 388 CGAGCCACCTGCCCCCGCCCGCCGCTACCACTCTGGGAAACCTCTTCGACGTA 447
Db 22 CGAGGAAATCTCTCTGCCCCCTCTCCAGTGATGGAATATCTTACAGATA 81
QY 448 CGCCCCGGGCGCTGTATTTCAGGCTCATGCGGCTAGTAAGAAAGTACGACCGGTGTC 507
Db 82 GGTATTAAAGACATCAGCAATCTTAACTCTCAAGGCTATGCCCCGTGTTTC 141
QY 508 ACCATCTACTGGGACCTGGCGGCTGTGGTCTGTGGTCTGGGAGGAGGTGTCGG 567
Db 142 ACTCTGTATTTGGCCCT--GAAACCCATAGTGGTCTGCATGATATGAAGCAGTGAAG 198
QY 568 GAGGCCCTGGAGGTTCAGCTGAGGAGTTCAGCGGCGGGGACCCGTAGCATGCTGGA 627
Db 199 GAAGCCCTGATTGATCTTGGAGAGAGGATTTCTGGAAGAGGCAATTTTCCACATGGGTGAA 258
QY 628 GGACATTTTATGATGGCCATGGGGTTTTCTTCTCAACGGGGAGCGGTGGAGCGAGCTGAGG 687
Db 259 AGAGCTAACAGAGGATTTGGAATTTTTCAGCAATGGAAGAAATGGAAGGAGATCCGG 318
QY 688 AGTTTACCATGCTTGTCTGCGGGACCTGTGGGATGGGGAAGCGAGAGGAGCTG 747
Db 319 CGTTTCTCCTCATGACGCTCGGAATTTTGGGATGGGGAAGAGGAGCATTTAGGACCGT 378
QY 748 ATCCAGCGGAGCCCGGTGTCTGTGGGAGACATTCAGGCGGACAGAGACCCCATTC 807
Db 379 GTTCAAGAGAGAGCCCGCTGCTGTGGAGGAGTTGAGAAAAACCAAGGCTTCCACCTGT 438
QY 808 GATCCCTCTCTGCTGTCGCCAGCCACCTCCACAGCTAGTCTGCTCCCTCTCTTGGC 867
Db 439 GATCCCACTTTTATCTCTGGGCTGTCTCTGCAATGTGATCTGCTCCATTTATTTCCAT 498
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 17, 2004, 14:58:20 ; Search time 1305.36 Seconds  
(without alignments)  
11383.562 Million cell updates/sec

Title: US-10-669-693-1

Perfect score: 2944

Sequence: 1 tttctctgttcttactc.....aaaaaaaaaaaaaaaaa 2944

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3327077 seqs, 2523723180 residues

Total number of hits satisfying chosen parameters: 6654154

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:\*

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*
- 12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*
- 13: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*
- 14: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*
- 15: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*
- 16: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*
- 17: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*
- 18: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*
- 19: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	2944	100.0	2944	9	US-09-748-127-1
2	2944	100.0	2944	13	US-10-669-693-1
3	2621.4	89.0	2623	13	US-10-147-493-63
4	2621.4	89.0	2623	13	US-10-145-127-63
5	2621.4	89.0	2623	13	US-10-160-503-63
6	2621.4	89.0	2623	13	US-10-143-118-63
7	2621.4	89.0	2623	13	US-10-144-993-63
8	2621.4	89.0	2623	13	US-10-158-787-63
9	2621.4	89.0	2623	13	US-10-140-024-63
10	2621.4	89.0	2623	13	US-10-140-808-63
11	2621.4	89.0	2623	13	US-10-152-405-63
12	2621.4	89.0	2623	13	US-10-127-852A-63
13	2621.4	89.0	2623	13	US-10-127-900A-63
14	2621.4	89.0	2623	13	US-10-128-685A-63

15	2621.4	89.0	2623	13	US-10-131-820A-63	Sequence 63, Appl
16	2621.4	89.0	2623	13	US-10-142-886-63	Sequence 63, Appl
17	2621.4	89.0	2623	13	US-10-146-728-63	Sequence 63, Appl
18	2621.4	89.0	2623	13	US-10-146-786-63	Sequence 63, Appl
19	2621.4	89.0	2623	13	US-10-147-499-63	Sequence 63, Appl
20	2621.4	89.0	2623	13	US-10-157-798-63	Sequence 63, Appl
21	2621.4	89.0	2623	15	US-10-028-072-63	Sequence 63, Appl
22	2621.4	89.0	2623	15	US-10-121-049-63	Sequence 63, Appl
23	2621.4	89.0	2623	15	US-10-123-904-63	Sequence 63, Appl
24	2621.4	89.0	2623	15	US-10-140-470-63	Sequence 63, Appl
25	2621.4	89.0	2623	15	US-10-175-746-63	Sequence 63, Appl
26	2621.4	89.0	2623	15	US-10-176-918-63	Sequence 63, Appl
27	2621.4	89.0	2623	15	US-10-176-921-63	Sequence 63, Appl
28	2621.4	89.0	2623	15	US-10-137-865-63	Sequence 63, Appl
29	2621.4	89.0	2623	15	US-10-140-474-63	Sequence 63, Appl
30	2621.4	89.0	2623	15	US-10-142-431-63	Sequence 63, Appl
31	2621.4	89.0	2623	15	US-10-143-114-63	Sequence 63, Appl
32	2621.4	89.0	2623	15	US-10-140-002-63	Sequence 63, Appl
33	2621.4	89.0	2623	15	US-10-142-419-63	Sequence 63, Appl
34	2621.4	89.0	2623	15	US-10-123-262-63	Sequence 63, Appl
35	2621.4	89.0	2623	15	US-10-123-262-63	Sequence 63, Appl
36	2621.4	89.0	2623	15	US-10-121-050-63	Sequence 63, Appl
37	2621.4	89.0	2623	15	US-10-141-755-63	Sequence 63, Appl
38	2621.4	89.0	2623	15	US-10-143-032-63	Sequence 63, Appl
39	2621.4	89.0	2623	15	US-10-123-108-63	Sequence 63, Appl
40	2621.4	89.0	2623	15	US-10-123-236-63	Sequence 63, Appl
41	2621.4	89.0	2623	15	US-10-123-261-63	Sequence 63, Appl
42	2621.4	89.0	2623	15	US-10-140-921-63	Sequence 63, Appl
43	2621.4	89.0	2623	15	US-10-140-928-63	Sequence 63, Appl
44	2621.4	89.0	2623	15	US-10-121-045-63	Sequence 63, Appl
45	2621.4	89.0	2623	15	US-10-123-292-63	Sequence 63, Appl

ALIGNMENTS

RESULT 1  
US-09-748-127-1  
; Sequence 1, Application US/09748127  
; Patent No. US20020076774A1  
; GENERAL INFORMATION:  
; APPLICANT: YAN, Chunhua et al  
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING  
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN  
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS,  
; TITLE OF INVENTION: AND USES THEREOF  
; FILE REFERENCE: CL000695  
; CURRENT APPLICATION NUMBER: US/09/748,127  
; CURRENT FILING DATE: 2001-06-11  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 2944  
; TYPE: DNA  
; ORGANISM: Human  
US-09-748-127-1

Query Match	100.0%;	Score 2944;	DB 9;	Length 2944;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 2944;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	TTTCTCTGTTGCTTACTCTCTATCCGGGGGCCCAAGCGGTGTCCTCGCGCCCAAGC	60	
DB	1	TTTCTCTGTTGCTTACTCTCTATCCGGGGGCCCAAGCGGTGTCCTCGCGCCCAAGC	60	
QY	61	CCCGGGTAACTGGTGACCTCCGAGACATCCGTTGAGCATGAGTTCCCGACATCAGG	120	
DB	61	CCCGGGTAACTGGTGACCTCCGAGACATCCGTTGAGCATGAGTTCCCGACATCAGG	120	
QY	121	CGGCGCGGTGGTCCGGGAGAAAACCCGCGGGGGAGATAGCCTGCCAGGCGCAGG	180	
DB	121	CGGCGCGGTGGTCCGGGAGAAAACCCGCGGGGGAGATAGCCTGCCAGGCGCAGG	180	























961	Db	 CAGCACACCGTCCGGGTATACCCCTCTGTCTCTGTATGAATAACCTCATGTCAAAAGTG	1020
1284	Qy	 GCTACGTGAGGAGCTGAATCGGAGCTGGGGGTGGCCAGGCACCAAGCTTAGGGGACCG	1343
1021	Db	 GGTACGTGAGGAGCTGATCGGAGCTGGGGGTGGCCAGGCACCAAGCTTAGGGGACCG	1080
1344	Qy	 TACCGGCTCCCTTTACACCGACGCGTTCTGTGANTAGGCGGAGCGGCTGTGGCGCTGGT	1403
1081	Db	 TACCGGCTCCCTTTACACCGACGCGTTCTGTGANTAGGCGGAGCGGCTGTGGCGCTGGT	1140
1404	Qy	 GCCCATGGGAATACCCCGACACCTCATGCGGACCAACCGTTCCGAGGGGTACACCTCGCC	1463
1141	Db	 GCCCATGGGAATACCCCGACACCTCATGCGGACCAACCGTTCCGAGGGGTACACCTCGCC	1200
1464	Qy	 CCAGGGCACGGAGTCTTTCCCGCTCTCTTAGGGAAGCGTGTCTGCCTTGGAGAGGGCTGGCAAA	1523
1201	Db	 CCAGGGCACGGAGTCTTTCCCGCTCTCTTAGGGAAGCGTGTCTGCCTTGGAGAGGGCTGGCAAA	1260
1524	Qy	 GCACCCAGAGAGTTCAACCCAGACCGTTTCTCTGGATGCAGATGAGACGGTTCAGGAAGCA	1583
1261	Db	 GCACCCAGAGAGTTCAACCCAGACCGTTTCTCTGGATGCAGATGAGACGGTTCAGGAAGCA	1320
1584	Qy	 TGAGGCGTTCCTGCGCTCTCTCTTAGGGAAGCGTGTCTGCCTTGGAGAGGGCTGGCAAA	1643
1321	Db	 TGAGGCGTTCCTGCGCTCTCTCTTAGGGAAGCGTGTCTGCCTTGGAGAGGGCTGGCAAA	1380
1644	Qy	 AGCGGAGTCTTCTCTTTCTTCCACCACTCTCTACAGCGTTCTCTCCTGGAGAGCCGTG	1703
1391	Db	 AGCGGAGTCTTCTCTTTCTTCCACCACTCTCTACAGCGTTCTCTCCTGGAGAGCCGTG	1440
1704	Qy	 CCCGCCGAGCACCTTGAGCTCAAGCCGACCGTTCAGTGGCGTTTTCACCATTTCCCGCAGC	1763
1441	Db	 CCCGCCGAGCACCTTGAGCTCAAGCCGACCGTTCAGTGGCGTTTTCACCATTTCCCGCAGC	1500
1764	Qy	 CTTCCAGCTGCAAGTTCGGTCCACTGTACCTTCACTCCACCGACGACACAGATGAAGGAA	1823
1501	Db	 CTTCCAGCTGCAAGTTCGGTCCACTGTACCTTCACTCCACCGACGACACAGATGAAGGAA	1560
1824	Qy	 GGCACTTTGGAAGTGGTGGGTGCCGAGGAGGTGCCCTCCAGCGCTCAACAGTGGGCGATGGA	1883
1561	Db	 GGCACTTTGGAAGTGGTGGGTGCCGAGGAGGTGCCCTCCAGCGCTCAACAGTGGGCGATGGA	1620
1884	Qy	 CAGGGTTAATGTCTCCAGAGTGTACATGACAGGACGACCATTTACACGCGCTTCAGTTGT	1943
1621	Db	 CAGGGTTAATGTCTCCAGAGTGTACATGACAGGACGACCATTTACACGCGCTTCAGTTGT	1680
1944	Qy	 TTTTCCGGAGTCTGTCCACGGGCCACAGCGCTCACTTGACTCATGCTGCTTAAGATGCACAA	2003
1681	Db	 TTTTCCGGAGTCTGTCTCCACGGGCCACAGCGCTCACTTGACTCATGCTGCTTAAGATGCACAA	1740
2004	Qy	 CCGCACACCCATACAACTACAGGGGCCAATAAGAACTGTCTGGGTTAGTTTCCACAG	2063
1741	Db	 CCGCACACCCATACAACTACAGGGGCCAATAAGAACTGTCTGGGTTAGTTTCCACAG	1800
2064	Qy	 ACATAAATATAGTTCATCTGGAATCACAGCACATAGCCAGGTAACCCACCACTCCCGCT	2123
1801	Db	 ACATAAATATAGTTCATCTGGAATCACAGCACATAGCCAGGTAACCCACCACTCCCGCT	1860
2124	Qy	 GGATCTGACGCCACACAGTGGGAGTCTGGCTGTCACTTCAACAGCCACAGAAACGGCCA	2183
1861	Db	 GGATCTGACGCCACACAGTGGGAGTCTGGCTGTCACTTCAACAGCCACAGAAACGGCCA	1920
2184	Qy	 CACATGTTACAGCTCACAGCCCTCTCGAATTCATCGAACTTCTCAGTGTCCCTGTCCCT	2243
1921	Db	 CACATGTTACAGCTCACAGCCCTCTCGAATTCATCGAACTTCTCAGTGTCCCTGTCCCT	1980
2244	Qy	 GGTGCTGGCACAGGGAAACAGATGCCCGCTCCGGGTCATGCCACCCACAGAGACTGTGCG	2303
1981	Db	 GGTGCTGGCACAGGGAAACAGATGCCCGCTCCGGGTCATGCCACCCACAGAGACTGTGCG	2040
2304	Qy	 TGTCTATGGCCCCCAACTCATGCTTCCCTCTCTTGGTCAACACACTTCCACGCGCTGTGACC	2363

RESULT 7  
US-10-144-993-63  
: Sequence 63. Application US/10144993

```

/ GENERAL INFORMATION:
/ APPLICANT: Baker, Kevin P.
/ APPLICANT: Beresini, Maureen
/ APPLICANT: DeForge, Laura
/ APPLICANT: Desnoyers, Luc
/ APPLICANT: Filvaroff, Ellen
/ APPLICANT: Gao, Wei-Qiang
/ APPLICANT: Gerritsen, Mary E
/ APPLICANT: Goddard, Audrey
/ APPLICANT: Godowski, Paul J.
/ APPLICANT: Gurney, Austin L.
/ APPLICANT: Sherwood, Steven
/ APPLICANT: Smith, Victoria
/ APPLICANT: Stewart, Timothy A.
/ APPLICANT: Tamas, Daniel
/ APPLICANT: Watanabe, Colin K
/ APPLICANT: Wood, William
/ APPLICANT: Zhang, Zemin
/ TITLE OF INVENTION: SECRETED A
/ TITLE OF INVENTION: ACIDS EN
/ FILE REFERENCE: P330R1C261
/ CURRENT APPLICATION NUMBER: U
/ CURRENT FILING DATE: 2002-05
/ Prior Application removed - S
/ NUMBER OF SEQ ID NOS: 550
/ SEQ ID NO 63
/ LENGTH: 2623
/ TYPE: DNA
/ ORGANISM: Homo Sapien
/ 10-144-993-63

```

## RESULT 7

US-10-144-993-63  
; Sequence 63, Application US/10144993  
; Publication No. US20040038336A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: DeForge, Laura  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Sherwood, Steven  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; TITLE OF INVENTION: ACIDS ENCODING THE SAME

; TITLE OF INVENTION: ACIDS  
; FILE REFERENCE: P3330R1C261

FILE REFERENCE: F3330KIC261  
CURRENT APPLICATION NUMBER: US/10/144,993

: CURRENT FILING DATE: 2002-05-13  
 : CURRENT AFFIDAVIT NUMBER: 007-107-117-113

;; Prior Application removed

NUMBER OF SEQ ID NOS: 550

NUMBER OF CHS  
; SEQ ID NO 63

; LENGTH: 20

TYPE: DNA

ORGANISM: HO

US-10-144-993-63

; APPLICANT: Zhang, Zemin



Db	2101	ACCGATGCCACACACCCCAACCACTGTGCCACACAGCTACCCACGCTACAACATCGTCC	2160
Qy	2424	TGGTCCCGCAGAGTATCTTCCCACTGAGACACGCGGCCCCACACAGAGCAGTCCCGCAG	2483
Db	2161	TGGTCCCGCAGAGTATCTTCCCACTGAGACACGCGGCCCCACACAGAGCAGTCCCGCAG	2220
Qy	2484	CCACTCTGCAACTGCGAGCCCTCAGTCAACCCCTTTTAAAGCACCCCTGATTTCAAAATG	2543
Db	2221	CCACTCTGCAACTGCGAGCCCTCAGTCAACCCCTTTTAAAGCACCCCTGATTTCAAAATG	2280
Qy	2544	CAAAACATCTGGGTCTCGGATTTAGCACAGAGCTTTGGACATACGAGGACCTTCAGAC	2603
Db	2281	CAAAACATCTGGGTCTCGGATTTAGCACAGAGCTTTGGACATACGAGGACCTTCAGAC	2340
Qy	2604	CGGAGGACACTGCCCAACCCCAACACGTCGTATGTAACACAGCTGGAAGCGGCCCT	2663
Db	2341	CGGAGGACACTGCCCAACCCCAACACGTCGTATGTAACACAGCTGGAAGCGGCCCT	2400
Qy	2664	GCTGCCCTCCACACACACATACACATCACTGATCTACAGCCCTGTTCGGCGTCAGAG	2723
Db	2401	GCTGCCCTCCACACACACATACACATCACTGATCTACAGCCCTGTTCGGCGTCAGAG	2460
Qy	2724	TCCCACATAGACCAGTGGAGGGTTAGAGACCAAGTAGGGGCCAGTTTCCAAATTCACC	2783
Db	2461	TCCCACATAGACCAGTGGAGGGTTAGAGACCAAGTAGGGGCCAGTTTCCAAATTCACC	2520
Qy	2784	CTGTGAGGAGTGAGCCGATCTGACGTTCTTGTGACTTAAGGGTCCGGCTTGGGAATT	2843
Db	2521	CTGTGAGGAGTGAGCCGATCTGACGTTCTTGTGACTTAAGGGTCCGGCTTGGGAATT	2580
Qy	2844	AAAGTTTGTCTGCGCTTACCTTAAACCAAAAAAAAAAAAAA	2886
Db	2581	AAAGTTTGTCTGCGCTTACCTTAAACCAAAAAAAAAAAAAA	2623
RESULT 8			
US-10-158-787-63			
; Sequence 63, Application US/10158787			
; Publication No. US20040039164A1			
; GENERAL INFORMATION:			
; APPLICANT: Baker, Kevin P.			
; APPLICANT: Beresini, Maureen			
; APPLICANT: DeForge, Laura			
; APPLICANT: Desnoyers, Luc			
; APPLICANT: Filvaroff, Ellen			
; APPLICANT: Gao, Wei-Qiang			
; APPLICANT: Gerritsen, Mary E.			
; APPLICANT: Goddard, Audrey			
; APPLICANT: Godowski, Paul J.			
; APPLICANT: Gurney, Austin L.			
; APPLICANT: Sherwood, Steven			
; APPLICANT: Smith, Victoria			
; APPLICANT: Stewart, Timothy A.			
; APPLICANT: Tumas, Daniel			
; APPLICANT: Watanabe, Colin K			
; APPLICANT: Wood, William			
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC			
; TITLE OF INVENTION: ACIDS ENCODING THE SAME			
; FILE REFERENCE: P3330R1C449			
; CURRENT APPLICATION NUMBER: US/10/158,787			
; CURRENT FILING DATE: 2003-04-03			
; PRIOR APPLICATION NUMBER: 60/049911			
; PRIOR FILING DATE: 1997-06-18			
; PRIOR APPLICATION NUMBER: 60/056974			
; PRIOR FILING DATE: 1997-08-26			
; PRIOR APPLICATION NUMBER: 60/059113			
; PRIOR FILING DATE: 1997-09-17			
; PRIOR APPLICATION NUMBER: 60/059115			
; PRIOR FILING DATE: 1997-09-17			
; PRIOR APPLICATION NUMBER: 60/059117			
; PRIOR FILING DATE: 1997-09-17			

; PRIOR APPLICATION NUMBER: 60/059122			
; PRIOR FILING DATE: 1997-09-17			
; PRIOR APPLICATION NUMBER: 60/059184			
; PRIOR FILING DATE: 1997-09-17			
; PRIOR APPLICATION NUMBER: 60/059263			
; PRIOR FILING DATE: 1997-09-18			
; PRIOR APPLICATION NUMBER: 60/059352			
; PRIOR FILING DATE: 1997-09-19			
; PRIOR APPLICATION NUMBER: 60/059588			
; PRIOR FILING DATE: 1997-09-19			
; Remaining Prior Application data removed - See File Wrapper or PALM.			
; NUMBER OF SEQ ID NOS: 550			
; SEQ ID NO 63			
; LENGTH: 2623			
; TYPE: DNA			
; ORGANISM: Homo Sapien			
US-10-158-787-63			
Query Match			
Best Local Similarity 89.0%; Score 2621.4; DB 13; Length 2623;			
Matches 2622; Conservative 0; Mismatches 1; Indels 0; Gaps 0;			
Qy	264	CGGAGCGCCCTGGGAGGAGGAGGAGCGGACCTGCGAGATGGAGCGACCGGACCTG 323	
Db	1	CGGAGCGCCCTGGGAGGAGGAGGAGCGGACCTGCGAGATGGAGCGACCGGACCTG 60	
Qy	324	GGCGTGTCTGGCGCTGGCGCTGCTCTCTGTGTGACGCTGGCGCTGTCCGGGACCGAG 383	
Db	61	GGCGTGTCTGGCGCTGGCGCTGCTCTCTGTGTGACGCTGGCGCTGTCCGGGACCGAG 120	
Qy	384	GGCCCGAGGCGCACCTGCCCGCCCGGCGCCACCGCGCTACACTGTCTGGGAAACCTCTGCA 443	
Db	121	GGCCCGAGGCGCACCTGCCCGCCCGGCGCCACCGCGCTACACTGTCTGGGAAACCTCTGCA 180	
Qy	444	GCTACGGCCCGGGCGCTGTATTACGGGCTCATGCGGCTGAGTAGAGTAGCGACCGGT 503	
Db	181	GCTACGGCCCGGGCGCTGTATTACGGGCTCATGCGGCTGAGTAGAGTAGCGACCGGT 240	
Qy	504	GTTTACCACTTACCTGGGACCCCTGGCGGCTGTGTGTGTCTCTGGTGGGACGAGGCTGT 563	
Db	241	GTTTACCACTTACCTGGGACCCCTGGCGGCTGTGTGTGTCTCTGGTGGGACGAGGCTGT 300	
Qy	564	GGGGAGGCGCCCTGGGAGGCTCAGGCTGAGGAGTTTCAAGCGCCCGGGGAAACCGTAGCGATGCT 623	
Db	301	GGGGAGGCGCCCTGGGAGGCTCAGGCTGAGGAGTTTCAAGCGCCCGGGGAAACCGTAGCGATGCT 360	
Qy	624	GGAGGAGCTTTTGTATGGCCATGGGGTTTTCTTCTCCACGGGAGCGGCTGGAGGCGAGCT 683	
Db	361	GGAGGAGCTTTTGTATGGCCATGGGGTTTTCTTCTCCACGGGAGCGGCTGGAGGCGAGCT 420	
Qy	684	GAGGAAGTTTACCATGCTTGTCTCTGGGACCTGGGACATGGGACATGGGACGAGGAGGA 743	
Db	421	GAGGAAGTTTACCATGCTTGTCTCTGGGACCTGGGACATGGGACATGGGACGAGGAGGA 480	
Qy	744	GCTGATCCAGGCGGAGGCGCGGTGTCTGTGTGAGACATTTCCAGGGACACAGAGGACGCCC 803	
Db	481	GCTGATCCAGGCGGAGGCGCGGTGTCTGTGTGAGACATTTCCAGGGACACAGAGGACGCCC 540	
Qy	804	ATTGATCTCCCTCCCTGCTGTGCGCCAGGCGACCTCCAAAGTAGTCTGTCTCCCTCTCTT 863	
Db	541	ATTGATCTCCCTCCCTGCTGTGCGCCAGGCGACCTCCAAAGTAGTCTGTCTCCCTCTCTT 600	
Qy	864	TGGCCTCCGCTTCTCTTATGAGGATTAAGGAGTTTCCAGGGCGGTGGTCCGGGACGCTGGTGG 923	
Db	601	TGGCCTCCGCTTCTCTTATGAGGATTAAGGAGTTTCCAGGGCGGTGGTCCGGGACGCTGGTGG 660	
Qy	924	TACCTGTGTGGAGTCACTCCAGGGGGTCACTACGAGATCTCTCTCTGTTTCTT 983	
Db	661	TACCTGTGTGGAGTCACTCCAGGGGGTCACTACGAGATCTCTCTCTGTTTCTT 720	
Qy	984	GGGGCCCTTCCAGGCGCCCAAGAGCTCTCTCCACACGCTCAGCACCTTGGCTGCTT 1043	
Db	721	GGGGCCCTTCCAGGCGCCCAAGAGCTCTCTCCACACGCTCAGCACCTTGGCTGCTT 780	

QY 1044 CACAGTCCGGCAGGTGACAGCAGCAGCGGAACTTGATGCTTTCGGGCGCCGACACGTGA 1103  
Db 781 CACAGTCCGGCAGGTGACAGCAGCAGCGGAACTTGATGCTTTCGGGCGCCGACACGTGA 840  
QY 1104 CTTTGTGATGCTTCTCTGCTGAAGATGGCA CAGGAGGAAACAAACCCAGGCA CAGAAT 1163  
Db 841 CTTTGTGATGCTTCTCTGCTGAAGATGGCA CAGGAGGAAACAAACCCAGGCA CAGAAT 900  
QY 1164 CACCAACAAGAACATGCTGATGACAGTCATTTATTTGCTTTGCTTGGGACGATGA CGGT 1223  
Db 901 CACCAACAAGAACATGCTGATGACAGTCATTTATTTGCTTTGCTTGGGACGATGA CGGT 960  
QY 1224 CAGCACCA CGGTCCGGCTATACCTCTCTGCTCTGATGAAATACCTCATGTCCRAAAGTG 1283  
Db 961 CAGCACCA CGGTCCGGCTATACCTCTCTGCTCTGATGAAATACCTCATGTCCRAAAGTG 1020  
QY 1284 GGTACGTGAGAGTGAAATCGGGAGCTGGGGCTGGCCAGGCA CCAAGCCTAGGGGACCG 1343  
Db 1021 GGTACGTGAGAGTGAAATCGGGAGCTGGGGCTGGCCAGGCA CCAAGCCTAGGGGACCG 1080  
QY 1344 TACCGCCTCCCTTACACCGACGGGTTCTGATGAGCGGAGCGGCTGTGGCGTGGT 1403  
Db 1081 TACCGCCTCCCTTACACCGACGGGTTCTGATGAGCGGAGCGGCTGTGGCGTGGT 1140  
QY 1404 GCCCATGGGAATACCCCGCACCTCTATGCGGACCA CCGCTTCGAGGGGTACACCCCTGCC 1463  
Db 1141 GCCCATGGGAATACCCCGCACCTCTATGCGGACCA CCGCTTCGAGGGGTACACCCCTGCC 1200  
QY 1464 CAGGGCA CGGAGGTCTTCCCTCTCTGCTCCATCTGATGACCCCAACATCTTCAA 1523  
Db 1201 CAGGGCA CGGAGGTCTTCCCTCTCTGCTCCATCTGATGACCCCAACATCTTCAA 1260  
QY 1524 GCACCCAGAGAGTCTCAACCCAGACCGTTTCTGGATGAGATGGA CGGTTCAGGAAGCA 1583  
Db 1261 GCACCCAGAGAGTCTCAACCCAGACCGTTTCTGGATGAGATGGA CGGTTCAGGAAGCA 1320  
QY 1584 TGAGGGCTTCTCTCCCTCTCTAGGAAAGCTGTCTGCTTGGAGGGCGCTGGCAA 1643  
Db 1321 TGAGGGCTTCTCTCCCTCTCTAGGAAAGCTGTCTGCTTGGAGGGCGCTGGCAA 1380  
QY 1644 AGCGAGCTTCTCTCTTCTTCCACCACTCTCTACAGCCTTCTCTGAGAGCGCGTG 1703  
Db 1381 AGCGAGCTTCTCTCTTCTTCCACCACTCTCTACAGCCTTCTCTGAGAGCGCGTG 1440  
QY 1704 CCGCGCGACACCTCAGCTCAGCCCAACCGTTCAGTGGCTTTTCAAANTTCCCGCAGC 1763  
Db 1441 CCGCGCGACACCTCAGCTCAGCCCAACCGTTCAGTGGCTTTTCAAACATTTCCCGCAGC 1500  
QY 1764 CTTCCAGCTGCAAGTCCGCTCCCACTGACCTTCACTCCACCGACGACGATGAAGGA 1823  
Db 1501 CTTCCAGCTGCAAGTCCGCTCCCACTGACCTTCACTCCACCGACGACGATGAAGGA 1560  
QY 1824 GGCACCTTGGAGTGGTGGTGGCCAGGACGCTGCTCCAGCCTCAACAGTGGGCAATGA 1883  
Db 1561 GGCACCTTGGAGTGGTGGTGGCCAGGACGCTGCTCCAGCCTCAACAGTGGGCAATGA 1620  
QY 1884 CAGGGTTAATGCTCCAGAGTGTACACTCAGGAGCGGACCAATTACAGCCTGCGATGTTG 1943  
Db 1621 CAGGGTTAATGCTCCAGAGTGTACACTCAGGAGCGGACCAATTACAGCCTGCGATGTTG 1680  
QY 1944 TTTCCGGAGTCTGTCCACGGGCCACACGCTCACTTGTACTGCTGCTAAGATGCACAA 2003  
Db 1681 TTTCCGGAGTCTGTCCACGGGCCACACGCTCACTTGTACTGCTGCTAAGATGCACAA 1740  
QY 2004 CGGCA CACCCATACAACTA CAAAGGGCCACAAAGCAACTGCTGGGTAGCTTTCCACAG 2063  
Db 1741 CGGCA CACCCATACAACTA CAAAGGGCCACAAAGCAACTGCTGGGTAGCTTTCCACAG 1800  
QY 2064 ACATAAATATAGTCCATCTGCAATCA CAGAGCATA GCGAGGTAA CCAACCACTCCCT 2123  
Db 1801 ACATAAATATAGTCCATCTGCAATCA CAGAGCATA GCGAGGTAA CCAACCACTCCCT 1860

QY 2124 GGATCTGCAGCCCA CACGTGGAGTCTGGCTGTCA CTTTCA CAGGCCACAGAAACGGCCA 2183  
Db 1861 GGA TCTGCAGCCCA CACGTGGAGTCTGGCTGTCA CTTTCA CAGGCCACAGAAACGGCCA 1920  
QY 2184 CACATGTTACAGCTCACAGCCCTCTCCATTCATCGAACTTCTCAGTGTCCCTTCCCT 2243  
Db 1921 CACATGTTACAGCTCACAGCCCTCTCCATTCATCGAACTTCTCAGTGTCCCTTCCCT 1980  
QY 2244 GGTGCTGGCACAGGAA CAGCATGCCCCCTCCGGGTCA TGCCACCCAGAGACTGTGCG 2303  
Db 1981 GGTGCTGGCACAGGAA CAGCATGCCCCCTCCGGGTCA TGCCACCCAGAGACTGTGCG 2040  
QY 2304 TGTCTATGGCCCCAACTCATGTCTCTCTTGGCTACA CACTCTCTCCAGCCTGTGACC 2363  
Db 2041 TGTCTATGGCCCCAACTCATGTCTCTCTTGGCTACA CACTCTCTCCAGCCTGTGACC 2100  
QY 2364 ACCGATGTCACACACACCCCAACCTTGTCCACAGACTACCCAGTACGACATCTGCC 2423  
Db 2101 ACCGATGTCACACACACCCCAACCTTGTCCACAGACTACCCAGTACGACATCTGCC 2160  
QY 2424 TGGTCTCCCGAGAGTATCTTCCCACTGAGACACGCGGCCCCACACAGAGGCA CAGTCCCCAG 2483  
Db 2161 TGGTCTCCCGAGAGTATCTTCCCACTGAGACACGCGGCCCCACACAGAGGCA CAGTCCCCAG 2220  
QY 2484 CCACCTCTGCAACTGCAAGCCCTCAGTCACCCCTTTTAAAGACCCCTGATTTACCAAATG 2543  
Db 2221 CCACCTCTGCAACTGCAAGCCCTCAGTCACCCCTTTTAAAGACCCCTGATTTACCAAATG 2280  
QY 2544 CAAACACATCTGGTCTGCGATTATGACAGAGACTTTTGGACATACAGAGCCCTCAGAC 2603  
Db 2281 CAAACACATCTGGTCTGCGATTATGACAGAGACTTTTGGACATACAGAGCCCTCAGAC 2340  
QY 2604 CGAGGAAACCTGCCCCAACCCCAACACGCTGCTTATGTAACCACTGTAAGAGCGGCCCT 2663  
Db 2341 CGAGGAAACCTGCCCCAACCCCAACACGCTGCTTATGTAACCACTGTAAGAGCGGCCCT 2400  
QY 2664 GCTGCCCTCCACACACACATACACACTACTGATCTACAGCCCTGTTCCGGGTACAG 2723  
Db 2401 GCTGCCCTCCACACACACATACACACTACTGATCTACAGCCCTGTTCCGGGTACAG 2460  
QY 2724 TCCCACTAGACCCAGTGGAGGGGTAGAGACCAAGTAGGGGCCAGTTTCCAAATCACC 2783  
Db 2461 TCCCACTAGACCCAGTGGAGGGGTAGAGACCAAGTAGGGGCCAGTTTCCAAATCACC 2520  
QY 2784 CTGTAGGGAGTGGAGCGGATCTGACCTTCTTGTGACTTAAGGCTCCGCTTGGGAAT 2843  
Db 2521 CTGTAGGGAGTGGAGCGGATCTGACCTTCTTGTGACTTAAGGCTCCGCTTGGGAAT 2580  
QY 2844 AAAGTTTGTCTGGCTTTAGCTTAAAAA AAAAAAAAAA 2886  
Db 2581 AAAGTTTGTCTGGCTTTAGCTTAAAAA AAAAAAAAAA 2623

## RESULT 9

US-10-140-024-63

; Sequence 63, Application US/10140024

; Publication NO. US20040058424A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Beresini, Maureen

; APPLICANT: DeForge, Laura

; APPLICANT: Desnoyers, Luc

; APPLICANT: Filvaroff, Ellen

; APPLICANT: Gao, Wei-Qiang

; APPLICANT: Gerritsen, Mary E.

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Sherwood, Steven

; APPLICANT: Smith, Victoria

; APPLICANT: Stewart, Timothy A.

; APPLICANT: Tumas, Daniel

; APPLICANT: Watanabe, Colin K



:	APPLICANT:	Wood,William			
:	APPLICANT:	Zhang,Zemin			
:	TITLE OF INVENTION:	SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC			
:	TITLE OF INVENTION:	ACIDS ENCODING THE SAME			
:	FILE REFERENCE:	P3330R1C69			
:	CURRENT APPLICATION NUMBER:	US/10/140,024			
:	CURRENT FILING DATE:	2002-05-06			
:	Prior Application removed - See Palm or File Wrapper				
:	NUMBER OF SEQ ID NOS:	550			
:	SEQ ID NO	63			
:	LENGTH:	2623			
:	TYPE:	DNA			
:	ORGANISM:	Homo Sapien			
:	US-10-140-024-63				
<hr/>					
Query Match		89.0%;	Score	2621.4;	DB 13; Length 2623;
Best Local Similarity		100.0%;	Pred. No.	0;	
Matches 2622;		Conservative	0;	Mismatches	1; Indels 0; Gaps 0;
<hr/>					
Qy	264	GCGGAGCGCTGGGAGAGGAAGAGGACCGACTGCCGAGATGGAGGCAGCCGACCTG 323			
Db	1	GCGGAGCGCTGGGAGAGGAGAGGACCGACTGCCGAGATGGAGGCAGCCGACCTG 60			
<hr/>					
Qy	324	GGCGCTGCTCTGGCGCTGGCGTGTCTCTGTCTGTGACGTGTGGCGCTGTTCGGGACCAG 383			
Db	61	GGCGCTGCTCTGGCGCTGGCGTGTCTCTGTCTGTGACGTGTGGCGCTGTTCGGGACCAG 120			
<hr/>					
Qy	384	GGCCCCGAGGCACCTGCCCCCGGGGCCCAAGCCGCTACCACTGCTGGGAAACCTCCTGCA 443			
Db	121	GGCCCCGAGGCACCTGCCCCCGGGGCCCAAGCCGCTACCACTGCTGGGAAACCTCCTGCA 180			
<hr/>					
Qy	444	GCTACGGCCC CGGGCGCTGTATTACGGGCTCATCGGGCTTGAGTAAGAAGTACGGACCGGT 503			
Db	181	GCTACGGCCC CGGGCGCTGTATTACGGGCTCATCGGGCTTGAGTAAGAAGTACGGACCGGT 240			
<hr/>					
Qy	504	GTTCCACCATCTACTTGGGAACCTTGGCGGCTGTGTGTGTCTTGGTGGGACGAGAGCTGT 563			
Db	241	GTTCCACCATCTACTTGGGAGACCTTGGCGGCGCTGTGTGTGTCTTGGTGGGACGAGAGCTGT 300			
<hr/>					
Qy	564	GCGGGAGGCCCTGGGAGGTACAGGCTGAGGAGTTCAGCGGCCGGGGAAACCTAGCGATGCT 623			
Db	301	GCGGGAGGCCCTGGGAGGTACAGGCTGAGGAGTTCAGCGGCCGGGGAAACCTAGCGATGCT 360			
<hr/>					
Qy	624	GGAGGAGCATTTTGATGCCCATGGGTTTTTCTTCTCAAACGGGAGCGGTGGAGGACGCT 683			
Db	361	GGAGGAGCATTTTGATGCCCATGGGTTTTTCTTCTCAAACGGGAGCGGTGGAGGACGCT 420			
<hr/>					
Qy	684	GAGGAAGTTTACCATGCTTGTCTGGGACCTGGGATGGGAGCGAAGCGCCAGGA 743			
Db	421	GAGGAAGTTTACCATGCTTGTCTGGGACCTGGGATGGGAGCGAAGCGCCAGGA 480			
<hr/>					
Qy	744	GCTGATCCAGCGGAGGCCCGGCTGTCTGTGGGAGACATTCACGGGGACAGAAAGACGCC 803			
Db	481	GCTGATCCAGCGGAGGCCCGGCTGTCTGTGGGAGACATTCACGGGGACAGAAAGACGCC 540			
<hr/>					
Qy	804	ATTGATTCCTCCCTGCTGTGGGCCACAGGCACCTCCAAAGTAGTCTGCTCCCTCCTCTT 863			
Db	541	ATTGATTCCTCCCTGCTGTGGGCCACAGGCACCTCCAAAGTAGTCTGCTCCCTCCTCTT 600			
<hr/>					
Qy	864	TGGCCTCGGCTTCTCTATAGGATAAGGAGTTCACGGCGGTGTCTCGGGCAGCTGGTGG 923			
Db	601	TGGCCTCGGCTTCTCTATAGGATAAGGAGTTCACGGCGGTGTCTCGGGCAGCTGGTGG 660			
<hr/>					
Qy	924	TACCTGTGGGAGTTCAGTCTCCACGGGGGGTTCAGACCTACAGAGTGTTCCTGGTTCCT 983			
Db	661	TACCTGTGGGAGTTCAGTCTCCACGGGGGGTTCAGACCTACAGAGTGTTCCTGGTTCCT 720			
<hr/>					
Qy	984	GCGGCCCTGACAGGCCCCACAGACGCTCCTCCACAGTACAGCACCTTGGCTGCCTT 1043			
Db	721	GCGGCCCTGACAGGCCCCACAGACGCTCCTCCACAGTACAGCACCTTGGCTGCCTT 780			
<hr/>					
Qy	1044	CACAGTCCGGCAGGTGCAGCAGCACAGGGGAACTGGATGCTTCGGGCCCGCACGTGA 1103			





Db 841 CCTGTGATGCTTCTGCTGAAGATGGCAACAGAGAAACAAACCCAGGCAAGAAATT 900  
QY 1164 CACCAACAAGAAATGCTGATGACAGTCAATTAATTTGCTGTTGCTGGAGCATGACGGT 1223  
Db 901 CACCAACAAGAAATGCTGATGACAGTCAATTAATTTGCTGTTGCTGGAGCATGACGGT 960  
QY 1224 CAGCAACACAGTGGCTATACCTGCTGCTCTGATGAATACCTCATGTCCTAAAGTG 1283  
Db 961 CAGCAACACAGTGGCTATACCTGCTGCTCTGATGAATACCTCATGTCCTAAAGTG 1020  
QY 1284 GGTACGTGAGAGCTGAATCGGGAGCTGGGGCTGGCCAGGCAACCAAGCTTAGGGACCG 1343  
Db 1021 GGTACGTGAGAGCTGAATCGGGAGCTGGGGCTGGCCAGGCAACCAAGCTTAGGGACCG 1080  
QY 1344 TACCGGCTCTCTTACACCGAGCGGTTCTGATGAGCGGAGCGGCTGCTGGCGCTGGT 1403  
Db 1081 TACCGGCTCTCTTACACCGAGCGGTTCTGATGAGCGGAGCGGCTGCTGGCGCTGGT 1140  
QY 1404 GCGCATGGGAATACCCCGGACCTCATGCGGACCAACCGCTTCCGAGGGTACACCTGCC 1463  
Db 1141 GCGCATGGGAATACCCCGGACCTCATGCGGACCAACCGCTTCCGAGGGTACACCTGCC 1200  
QY 1464 CAGGGCAAGAGTTCAACCCAGAGCGGTTCTGATGAGCGGAGCGGCTGCTGGCGCTGGT 1523  
Db 1201 CAGGGCAAGAGTTCAACCCAGAGCGGTTCTGATGAGCGGAGCGGCTGCTGGCGCTGGT 1260  
QY 1524 GCACCCAGAAGATTCAACCCAGAGCGGTTCTGATGAGCGGAGCGGTTGAGGAAGA 1583  
Db 1261 GCACCCAGAAGATTCAACCCAGAGCGGTTCTGATGAGCGGAGCGGTTGAGGAAGA 1320  
QY 1584 TGAGGGTTCTGCTTCTTACACCATCTTCAAGCGCTTCTCCCTGGAGAGCGCGTG 1703  
Db 1381 AGCGAGCTTCTTCTTCTTACACCATCTTCAAGCGCTTCTCCCTGGAGAGCGCGTG 1440  
QY 1704 CCGCGCGGACACCTGAGGCTCAAGCCCAACCGTCAAGTGGCTTTTCAACATTTCCCGCAGC 1763  
Db 1441 CCGCGCGGACACCTGAGGCTCAAGCCCAACCGTCAAGTGGCTTTTCAACATTTCCCGCAGC 1500  
QY 1764 CTTCCAGCTGCAAGTCCGTCCTCACTCACTTCACTCCACGACGACAGATGAAGAA 1823  
Db 1501 CTTCCAGCTGCAAGTCCGTCCTCACTCACTTCACTCCACGACGACAGATGAAGAA 1560  
QY 1824 GGCACCTTGAAGTGGTGGTGCCAGGACGCTGCTCCAGCCTCAACAGTGGGCAATGA 1883  
Db 1561 GGCACCTTGAAGTGGTGGTGCCAGGACGCTGCTCCAGCCTCAACAGTGGGCAATGA 1620  
QY 1884 CAGGGTTAATGCTCCAGAGTGTACACTGAGGAGCGCAATTTACAGCGCTGAGTTGT 1943  
Db 1621 CAGGGTTAATGCTCCAGAGTGTACACTGAGGAGCGCAATTTACAGCGCTGAGTTGT 1680  
QY 1944 TTTCGGAGTCTGCTCCACGGCCCAACGCTCACTGCTCATGCTGCTAAAGATGACAA 2003  
Db 1681 TTTCGGAGTCTGCTCCACGGCCCAACGCTCACTGCTCATGCTGCTAAAGATGACAA 1740  
QY 2004 CCGCACACCCATACAACTTCAAGGGCCCAAGCAACTGCTGGGTAGCTTTCCACAG 2063  
Db 1741 CCGCACACCCATACAACTTCAAGGGCCCAAGCAACTGCTGGGTAGCTTTCCACAG 1800  
QY 2064 ACATAAATATGCTCATCTGCAATCAACAGCATAGGCAAGTAAACCAACTCCCT 2123  
Db 1801 ACATAAATATGCTCATCTGCAATCAACAGCATAGGCAAGTAAACCAACTCCCT 1860  
QY 2124 GGATCTGACGCCACAGTGGAGTCTGGCTGCTCACTTCAAGGACACAGAAACGGCCA 2183  
Db 1861 GGATCTGACGCCACAGTGGAGTCTGGCTGCTCACTTCAAGGACACAGAAACGGCCA 1920  
QY 2184 CACATGTTACAGCTCAACGCCCTCTCCATTTCAATGCAATTTCTAGTGTCCCTGTCCCT 2243

Db 1921 CACATGTTACAGCTCAACGCCCTCTCCATTTATCGAACTTTCTAGTGTCCCTGTCCCT 1980  
QY 2244 GGTGCTTGGCACAGGAAACAGCATGCCCCCTTCCGGGGTTCATGCAACCCAGAGATGTGCGC 2303  
Db 1981 GGTGCTTGGCACAGGAAACAGCATGCCCCCTTCCGGGGTTCATGCAACCCAGAGATGTGCGC 2040  
QY 2304 TGTCTATGGCCCCAACTCATGCTCCTCTCTTGGGTACACCACTCTCCAGCTGTGAC 2363  
Db 2041 TGTCTATGGCCCCAACTCATGCTCCTCTCTTGGGTACACCACTCTCCAGCTGTGAC 2100  
QY 2364 ACCGATGTCACACACCCCAACCACTTGTCCACACAGCTACCCAGTACGACATGTGTC 2423  
Db 2101 ACCGATGTCACACACCCCAACCACTTGTCCACACAGCTACCCAGTACGACATGTGTC 2160  
QY 2424 TGGTCTCCCAAGTATCTTCCCACTGACACAGCGCGCCCAACAGAGGCAAGTCCCGAG 2483  
Db 2161 TGGTCTCCCAAGTATCTTCCCACTGACACAGCGCGCCCAACAGAGGCAAGTCCCGAG 2220  
QY 2484 CCACCTCTGCAACTGCGAGCCTCAGTCAACCTTTTAAAGCAACCTGATTTACCAATG 2543  
Db 2221 CCACCTCTGCAACTGCGAGCCTCAGTCAACCTTTTAAAGCAACCTGATTTACCAATG 2280  
QY 2544 CAAACATCTGGGCTCTGGATTTATGACAGAGACTTTGGACATACGAGGACCTCAGAC 2603  
Db 2281 CAAACATCTGGGCTCTGGATTTATGACAGAGACTTTGGACATACGAGGACCTCAGAC 2340  
QY 2604 CGGAGAAACACTGCGCCCAACCCCAACAGCTGTATGTAAACACAGTGGAAAGCGGCCCT 2663  
Db 2341 CGGAGAAACACTGCGCCCAACCCCAACAGCTGTATGTAAACACAGTGGAAAGCGGCCCT 2400  
QY 2664 GCTGCCCTTCCACACACATACACATCACTGATCTACAGCCCTGTTCCGCGTCAGAG 2723  
Db 2401 GCTGCCCTTCCACACACATACACATCACTGATCTACAGCCCTGTTCCGCGTCAGAG 2460  
QY 2724 TCCCACTAGACCCAGTGAAGGGTTAGAGACCAAGTAGAGGCGCAGTTTCCAAATTCACC 2783  
Db 2461 TCCCACTAGACCCAGTGAAGGGTTAGAGACCAAGTAGAGGCGCAGTTTCCAAATTCACC 2520  
QY 2784 CTGTAGGGAGTGAAGCGGATCTGAGCTTCTTGTGACTTAAGGGTCCGGCTTGGGAAT 2843  
Db 2521 CTGTAGGGAGTGAAGCGGATCTGAGCTTCTTGTGACTTAAGGGTCCGGCTTGGGAAT 2580  
QY 2844 AAAGTTGTTTCTGGCCTTTAGCCTTAAACCAACCAACCAACCAACCAACCAACCA 2886  
Db 2581 AAAGTTGTTTCTGGCCTTTAGCCTTAAACCAACCAACCAACCAACCAACCAACCA 2623

RESULT 11  
US-10-152-405-63  
; Sequence 63, Application US/10152405  
; Publication No. US20030211571A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: DeForge, Laura  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Sherwood, Steven  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; TITLE OF INVENTION: ACIDS ENCODING THE SAME  
; FILE REFERENCE: P3330R1C383  
; CURRENT APPLICATION NUMBER: US/10/152,405

```

; CURRENT FILING DATE: 2002-05-20
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 63
; LENGTH: 2623
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-152-405-63

```

Query Match		89.0%;	Score 2621.4;	DB 13;	Length 2623;
Best Local Similarity		100.0%;	Pred. No. 0;		
Matches 2622; Conservative		0;	Mismatches 1;	Indels	0; Gaps
Qy	264	GCGAGCGCCTGGGAGAGGAGGAGGAGCGGACCTGCCGAGATGGAGGCGACCGGACCTG	323		
Db	1	GCGAGCGCCTGGGAGAGGAGGAGGAGCGGACCTGCCGAGATGGAGGCGACCGGACCTG	60		
Qy	324	GGCGCTGCTGTGCGCTGGCGCTGCTCCTGCTGTGACGCTGGCGCTGTCGGGACCG	383		
Db	61	GGCGCTGCTGTGCGCTGGCGCTGCTCCTGCTGTGACGCTGGCGCTGTCGGGACCG	120		
Qy	384	GGCCCGAGGCCACTGCCCCCGGGCCCGACGCGGCTACCACTGTGGGAAACCTCCTGCA	443		
Db	121	GGCCCGAGGCCACTGCCCCCGGGCCCGACGCGGCTACCACTGTGGGAAACCTCCTGCA	180		
Qy	444	GCTACGGCCCCGGCGCTGTATTACAGGCTCATCGGCTGAGTAGAGTAGGACCGGT	503		
Db	181	GCTACGGCCCCGGCGCTGTATTACAGGCTCATCGGCTGAGTAGAGTAGGACCGGT	240		
Qy	504	GTTTCAACATCTACTGGGACCTGGCGCTGTGCTGCTGCTGGGACGAGGCTGT	563		
Db	241	GTTTCAACATCTACTGGGACCTGGCGCTGTGCTGCTGCTGGGACGAGGCTGT	300		
Qy	564	GCGGAGGCCCTGGGAGGTGAGGCTGAGGATTCAGCGCCGGGGAAACCGTAGCGATGCT	623		
Db	301	GCGGAGGCCCTGGGAGGTGAGGCTGAGGATTCAGCGCCGGGGAAACCGTAGCGATGCT	360		
Qy	624	GGAGGACTTTTGATGGCATGGGGTTTCTCTCCAAAGGGGAGCGGTGGAGGAGCT	683		
Db	361	GGAGGACTTTTGATGGCATGGGGTTTCTCTCCAAAGGGGAGCGGTGGAGGAGCT	420		
Qy	684	GAGGAGTTTACATGCTTGTGCGGACCTGGGATGGGAAAGCGAAGGCCGAGGA	743		
Db	421	GAGGAGTTTACATGCTTGTGCTGTGCGGACCTGGGATGGGAAAGCGAAGGCCGAGGA	480		
Qy	744	GCTGATCCAGCGGAGGCCCGGTGCTGCTGGAGACATTCAGGGGAGCAAGAGACGCC	803		
Db	481	GCTGATCCAGCGGAGGCCCGGTGCTGCTGGAGACATTCAGGGGAGCAAGAGACGCC	540		
Qy	804	ATTGATCCCTCCCTGCTGTGTCGCCAGGCCACCTCCAAAGTAGTGTGCTCCCTCTT	863		
Db	541	ATTGATCCCTCCCTGCTGTGTCGCCAGGCCACCTCCAAAGTAGTGTGCTCCCTCTT	600		
Qy	864	TGCGCTCCGCTTCTCTATGAGGATTAAGGATTCAGCGCGTGGTCCGGGACGTGGTG	923		
Db	601	TGCGCTCCGCTTCTCTATGAGGATTAAGGATTCAGCGCGTGGTCCGGGACGTGGTG	660		
Qy	924	TACCTGTCTGGGAGTCAGCTCCAGGGGCTCAGACCTACGAGATGTTCTCTGGTTCCT	983		
Db	661	TACCTGTCTGGGAGTCAGCTCCAGGGGCTCAGACCTACGAGATGTTCTCTGGTTCCT	720		
Qy	984	GGGGCCCTGCCAGGCCCCCAAGCAGCTCTCTCCACACGTCAGCACCTTGGCTGCCTT	1043		
Db	721	GGGGCCCTGCCAGGCCCCCAAGCAGCTCTCTCCACACGTCAGCACCTTGGCTGCCTT	780		
Qy	1044	CACAGTCCGGCAGGTGCAGGACCAAGGGAACTTGGATGTTTGGGCCCCGACGTGA	1103		
Db	781	CACAGTCCGGCAGGTGCAGGACCAAGGGAACTTGGATGTTTGGGCCCCGACGTGA	840		
Qy	1104	CCTTGTGATGCTTCTCTGTAAGATGTCACAGGAGGAACAAACCGACGACGAAT	1163		
Db	841	CCTTGTGATGCTTCTCTGTAAGATGTCACAGGAGGAACAAACCGACGACGAAT	900		

```

; PRIOR APPLICATION NUMBER: 60/056974
; PRIOR FILING DATE: 1997-08-26
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059115
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059117
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059122
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059184
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059588
; PRIOR FILING DATE: 1997-09-19
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 63
; LENGTH: 2623
; TYPE: DNA
; ORGANISM: Homo Sapien
; US-10-127-852A-63

Query Match      89.0%; Score 2621.4; DB 13; Length 2623;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 262; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      264  GCGGAGCGCTGGGAGGAGGAGGAGCGACCTGCCAGATGGAGCGCACGGCACCTG 323
Db      1    GCGGAGCGCTGGGAGGAGGAGGAGCGACCTGCCAGATGGAGCGCACGGCACCTG 60

Qy      324  GGCCTGTCTCTGGCGTGGCGCTGTCTTCGTCTGTACGCTGGCGCTGTCCGGGACCAG 383
Db      61    GGCCTGTCTCTGGCGTGGCGCTGTCTTCGTCTGTACGCTGGCGCTGTCCGGGACCAG 120

Qy      384  GGCCCGGAGGCACCTGTCCCCCGGGGCCACCGCCGCTACCACTGCTGGGAAACCTTCCTGCA 443
Db      121  GGCCCGGAGGCACCTGTCCCCCGGGGCCACCGCCGCTACCACTGCTGGGAAACCTTCCTGCA 180

; 444  GCTTCCGCGCGCGCGCTGTATTCACGGCTTCATGCGCTGCTAGTAAGAAGTACGGACCGGT 503

```

181	Db	GCTACGGCCCGGGCGCTGTATTACGGGCTCATCGGCLGATAGAGAGATACGACACGGT	241
504	QY	GTTCAACCATCTACCTGGGACCCCTGGCGGCTGTGGTGCTCTGTGTTGGCAGGAGGCTGT	563
241	Db	GTTCAACCATCTA	300
564	QY	CGCGGAGGCCCTGGGAGGTTCAGGTGAGGAGTTTCAGCGCCCGGGGAACCGTAGCGATGCT	623
301	Db	CGCGGAGGCCCTGGGAGGTTCAGGCTGAGGAGTTTCAGCGCCCGGGGAACCGTAGCGATGCT	360
624	QY	GGAGGGACATTTTCATGCGCAATGGGTTTTCTTCTCAAACGGGAGCGGTGGAGGACGCT	683
361	Db	GGAGGGACATTTTCATGCGCAATGGGTTTTCTTCTCAAACGGGAGCGGTGGAGGACGCT	420
684	QY	GAGGAAATTTACCATGCTTGCTCTCGGGACCTGGGCAATGGGGAAGCGAGAAGGCGAGGA	743
421	Db	GAGGAAATTTACCATGCTTGCTCTCGGGACCTGGGCAATGGGGAAGCGAGAAGGCGAGGA	480
744	QY	GCTGATCCAGCGGAGGCCCCGGTGTCTGGTGAGACATTCACAGGGGACAGAGAAGCGGCC	803
481	Db	GCTGATCCAGCGGAGGCCCCGGTGTCTGGTGAGACATTCACAGGGGACAGAGAAGCGGCC	540
804	QY	ATTGATCCCTTCCTGCTGCTGGCCAGGCCACCTTCCAAACGTAGTGTGCTCCCTCCTCTTT	863
541	Db	ATTGATCCCTTCCTGCTGCTGGCCAGGCCACCTTCCAAACGTAGTGTGCTCCCTCCTCTTT	600
864	QY	TGGCTCCGCTTCCTATAGGATAAGAGGTTTCAGGCCGTGGTCCGGGACGCTGGTGG	923
601	Db	TGGCTCCGCTTCCTATAGGATAAGAGGTTTCAGGCCGTGGTCCGGGACGCTGGTGG	660

QY 924 TACCTGCTGGAGTCACTCCAGGGGGTCCAGACCTACGAGATGTTCTCTGTTCT 983  
DB 661 TACCTGCTGGAGTCACTCCAGGGGGTCCAGACCTACGAGATGTTCTCTGTTCT 720  
QY 984 GGGGCCCCGAGGCCCCCAAGAGCTCTCCACCAAGTCCAGACCTTGGCTGCCTT 1043  
DB 721 GGGGCCCCGAGGCCCCCAAGAGCTCTCCACCAAGTCCAGACCTTGGCTGCCTT 780  
QY 1044 CACAGTCCGGAGGTCAGCAGCAGCCAGGGAACTGATGTTCTGGGCCCCGGAG 1103  
DB 781 CACAGTCCGGAGGTCAGCAGCAGCCAGGGAACTGATGTTCTGGGCCCCGGAG 840  
QY 1104 CTTGTGATGCTTCTGCTGAAGATGCAAGAGGAAACAAACCCAGGACAGAAAT 1163  
DB 841 CTTGTGATGCTTCTGCTGAAGATGCAAGAGGAAACAAACCCAGGACAGAAAT 900  
QY 1164 CACCAAGAAACATGCTGATGACAGTCAATTTATTTGTTGTTGCTGGGACGATGACGGT 1223  
DB 901 CACCAAGAAACATGCTGATGACAGTCAATTTATTTGTTGTTGCTGGGACGATGACGGT 960  
QY 1224 CAGCACACGGTCCGGCTATACCTCCCTGCTCTGATGAAATACCTCATGTTCCAAAGTG 1283  
DB 961 CAGCACACGGTCCGGCTATACCTCCCTGCTCTGATGAAATACCTCATGTTCCAAAGTG 1020  
QY 1284 GGTACGTGAGAGCTGAATCGGGAGCTGGGGGCTGGCCAGGACCAAGCCTAGGGACCG 1343  
DB 1021 GGTACGTGAGAGCTGAATCGGGAGCTGGGGGCTGGCCAGGACCAAGCCTAGGGACCG 1080  
QY 1344 TACCCGCTCCCTTACACCGACGGTCTGTCATGAGCGCAGGGCTGCTGGGCTGGT 1403  
DB 1081 TACCCGCTCCCTTACACCGACGGTCTGTCATGAGCGCAGGGCTGCTGGGCTGGT 1140  
QY 1404 GCCATGGGAATACCCCGACCTCATGCGGACCAAGCCTGCTCGAGGGTACACCTTGCC 1463  
DB 1141 GCCATGGGAATACCCCGACCTCATGCGGACCAAGCCTGCTCGAGGGTACACCTTGCC 1200  
QY 1464 CAGGGCAGCGAGTCTTCCCTCTTCCCTCTGCTGCTGATGACCCCAACATCTTCAA 1523  
DB 1201 CAGGGCAGCGAGTCTTCCCTCTTCCCTCTGCTGCTGATGACCCCAACATCTTCAA 1260  
QY 1524 GCACCCAGAGATTCACCCAGACCGTCTTCTGATGAGAGCGTTCAGGAAGCA 1583  
DB 1261 GCACCCAGAGATTCACCCAGACCGTCTTCTGATGAGAGCGTTCAGGAAGCA 1320  
QY 1584 TGAGGGCTTCTGCTTCTTCTAGGAGCGTCTGCTGCTGAGAGCGCTGGCAAA 1643  
DB 1321 TGAGGGCTTCTGCTTCTTCTAGGAGCGTCTGCTGCTGAGAGCGCTGGCAAA 1380  
QY 1644 AGCGAGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1703  
DB 1381 AGCGAGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1440  
QY 1704 CCGCGCGACACCTGAGCTCAAGCCACCGTCACTGAGCGCTTCTTCTTCTTCTTCT 1763  
DB 1441 CCGCGCGACACCTGAGCTCAAGCCACCGTCACTGAGCGCTTCTTCTTCTTCTTCT 1500  
QY 1764 CTTCCAGCTGCAAGTCCGTCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1823  
DB 1501 CTTCCAGCTGCAAGTCCGTCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1560  
QY 1824 GGCACCTTGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1883  
DB 1561 GGCACCTTGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1620  
QY 1884 CAGGGTAAATGCTCAGAGTGTACCTGAGCGACCAATTTACAGCCTGCGATGCT 1943  
DB 1621 CAGGGTAAATGCTCAGAGTGTACCTGAGCGACCAATTTACAGCCTGCGATGCT 1680  
QY 1944 TTTCCGGAGTCTGCCCGGCCCCACGCTCACTGATCTGCTGCTGCTGCTGCTGCTGCT 2003  
DB 1681 TTTCCGGAGTCTGCCCGGCCCCACGCTCACTGATCTGCTGCTGCTGCTGCTGCTGCT 1740

QY 2004 CCGGACACCCATACAACTACAAGGGCCACAAAGCAACTGCTGGGTAGCTTTCCACAG 2063  
DB 1741 CCGGACACCCATACAACTACAAGGGCCACAAAGCAACTGCTGGGTAGCTTTCCACAG 1800  
QY 2064 ACATAAATATAGTCTCATCTGCAATCAAGGACATAGGAGTAAACCCACCACTCCCT 2123  
DB 1801 ACATAAATATAGTCTCATCTGCAATCAAGGACATAGGAGTAAACCCACCACTCCCT 1860  
QY 2124 GGATCTGAGGCCACACCGTGGGAGTCTGGCTGTACCTTCAAGGCCACAGAAACGGCA 2183  
DB 1861 GGATCTGAGGCCACACCGTGGGAGTCTGGCTGTACCTTCAAGGCCACAGAAACGGCA 1920  
QY 2184 CACATGTTCAAGCTCAACGCCCTCTCAATCATTCGAACTTCTCAGTGTCCCTGCTCT 2243  
DB 1921 CACATGTTCAAGCTCAACGCCCTCTCAATCATTCGAACTTCTCAGTGTCCCTGCTCT 1980  
QY 2244 GGTGCTGGCACAGGACAGCATGCCCTCCCTCGGGTCAAGGCCACAGAGCTGTGCG 2303  
DB 1981 GGTGCTGGCACAGGACAGCATGCCCTCCCTCGGGTCAAGGCCACAGAGCTGTGCG 2040  
QY 2304 TGTCTATGGCCCCCAACTCATGCTCCCTCTCTTGGCTACACCACTCTCCAGCTGTGACC 2363  
DB 2041 TGTCTATGGCCCCCAACTCATGCTCCCTCTCTTGGCTACACCACTCTCCAGCTGTGACC 2100  
QY 2364 ACCGATGTCCACACACCCCCCAACCTTGTCCACACAGCTACCCACCTGATCGTCC 2423  
DB 2101 ACCGATGTCCACACACCCCCCAACCTTGTCCACACAGCTACCCACCTGATCGTCC 2160  
QY 2424 TGGCTCCCCAGAGTATCTCCACTGAGACACGCGCCCCACAGAGGACAGTCCCCAG 2483  
DB 2161 TGGCTCCCCAGAGTATCTCCACTGAGACACGCGCCCCACAGAGGACAGTCCCCAG 2220  
QY 2484 CCACCTCTGCAACTGACAGCCTCAGTCAACCCCTTTTAAAGCACCTGATTTCAAAATG 2543  
DB 2221 CCACCTCTGCAACTGACAGCCTCAGTCAACCCCTTTTAAAGCACCTGATTTCAAAATG 2280  
QY 2544 CAACACATCTGGTCTGCGATTATGACAGAGACTTTGGACATACAGGACCTCAGAC 2603  
DB 2281 CAACACATCTGGTCTGCGATTATGACAGAGACTTTGGACATACAGGACCTCAGAC 2340  
QY 2604 CGGAGGACACCTGCCCCAACCCCAACAGTGTCTTAAACCAAGTGGAAAGGCGCCCT 2663  
DB 2341 CGGAGGACACCTGCCCCAACCCCAACAGTGTCTTAAACCAAGTGGAAAGGCGCCCT 2400  
QY 2664 GCTGCCCTCCACACACATACATCACTCATCTGATCTACAGCCCTGTTCCGGGTCAGAG 2723  
DB 2401 GCTGCCCTCCACACACATACATCACTCATCTGATCTACAGCCCTGTTCCGGGTCAGAG 2460  
QY 2724 TCCCCACTAGACCCAGTGGAGGGTTAGAGCAAGTAGGGGCCAGTTTCCCAATTCACC 2783  
DB 2461 TCCCCACTAGACCCAGTGGAGGGTTAGAGCAAGTAGGGGCCAGTTTCCCAATTCACC 2520  
QY 2784 CTGTACGGAGTGAAGCGGATCTGAGTTCCTGTGACTTAAGGTCGCGCTGGGAAT 2843  
DB 2521 CTGTACGGAGTGAAGCGGATCTGAGTTCCTGTGACTTAAGGTCGCGCTGGGAAT 2580  
QY 2844 AAAGTTGTTTCTGGCCTTTAGCCTAAAAAATAAAAAA 2886  
DB 2581 AAAGTTGTTTCTGGCCTTTAGCCTAAAAAATAAAAAA 2623

RESULT 13  
US-10-127-900A-63  
; Sequence 63, Application US/10127900A  
; Publication No. US20030203429A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: DeForge, Laura  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerritsen, Mary E.



APPLICANT: Goddard, Audrey	Db	361	GGAAGGAGCTTTTGAATGGCCATGGGGTTTCTTCTCCAAAGGGAGCGGTGGAGGACGT	420
APPLICANT: Godowski, Paul J.	Qy	684	GAGGAAGTTTACCATGCTTGTCTGCGGGACCTCGGCATGGGGAGCGAGAGCGGAGGA	743
APPLICANT: Gurney, Austin L.	Db	421	GAGGAAGTTTACCATGCTTGTCTGCGGGACCTCGGCATGGGGAGCGAGAGCGGAGGA	480
APPLICANT: Smith, Victoria	Qy	744	GCTGATCCAGCGGAGGCGCGGTGTCTGTGGAGACATTCAGGGGACAGAGGACGCC	803
APPLICANT: Stewart, Timothy A.	Db	481	GCTGATCCAGCGGAGGCGCGGTGTCTGTGGAGACATTCAGGGGACAGAGGACGCC	540
APPLICANT: Tumas, Daniel	Qy	804	ATTGATCCCTCCCTGCTGCTGGCCAGGCCACCTCCAAAGCTAGTCTCTCTCCCTCTCT	863
APPLICANT: Watanabe, Colin K	Db	541	ATTGATCCCTCCCTGCTGCTGGCCAGGCCACCTCCAAAGCTAGTCTCTCTCCCTCTCT	600
APPLICANT: Wood, William	Qy	864	TGGCTCCGCTTCTCTATAGAGATAGAGAGTTCAGGCCGTGTCCGGGAGCTGTGGTGG	923
APPLICANT: Zhang, Zemin	Db	601	TGGCTCCGCTTCTCTATAGAGATAGAGAGTTCAGGCCGTGTCCGGGAGCTGTGGTGG	660
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC	Qy	924	TACCTGTGTGGAGTACGCTCCAGGGGGGTTCAGACCTACGAGATGTTCTCTCTGCTTCT	983
FILE REFERENCE: P330R1C81	Db	661	TACCTGTGTGGAGTACGCTCCAGGGGGGTTCAGACCTACGAGATGTTCTCTCTGCTTCT	720
CURRENT APPLICATION NUMBER: US/10/127,900A	Qy	984	GCGGCCCTGCCAGSCCCCAACAGCAGCTCTCTCCACCAAGCTCAGCAGCTTGGCTGCCTT	1043
CURRENT FILING DATE: 2002-10-15	Db	721	GCGGCCCTGCCAGSCCCCAACAGCAGCTCTCTCCACCAAGCTCAGCAGCTTGGCTGCCTT	780
PRIOR APPLICATION NUMBER: 60/049911	Qy	1044	CACAGTCCGGCAGGTGACAGCAGCACCAGGGGAACTGGATGTTTCGGGCCCCGACGTGA	1103
PRIOR FILING DATE: 1997-06-18	Db	781	CACAGTCCGGCAGGTGACAGCAGCACCAGGGGAACTGGATGTTTCGGGCCCCGACGTGA	840
PRIOR FILING DATE: 1997-08-26	Qy	1104	CCTTGTGATGCTTCTCTGCTGAAATGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGT	1163
PRIOR FILING DATE: 1997-09-17	Db	841	CCTTGTGATGCTTCTCTGCTGAAATGACAGGAGGAGGAGGAGGAGGAGGAGGAGT	900
PRIOR FILING DATE: 1997-09-17	Qy	1164	CACCAACAGAGACATGCTGATGACAGTCAATTTATTTGCTGTTGCTGGGACGATGACGCT	1223
PRIOR FILING DATE: 1997-09-17	Db	901	CACCAACAGAGACATGCTGATGACAGTCAATTTATTTGCTGTTGCTGGGACGATGACGCT	960
PRIOR FILING DATE: 1997-09-19	Qy	1224	CAGCACACGCTGCGGTATACCTCTCTGCTGATGAAATACCTCTATGTCCTCAAGTG	1283
PRIOR FILING DATE: 1997-09-19	Db	961	CAGCACACGCTGCGGTATACCTCTCTGCTGATGAAATACCTCTATGTCCTCAAGTG	1020
PRIOR FILING DATE: 1997-09-19	Qy	1284	GCTGATGAGGAGTGAATCGGGAGCTGGGGCTGGCCAGGACCAAGCTAGGGACCG	1343
PRIOR FILING DATE: 1997-09-19	Db	1021	GCTGATGAGGAGTGAATCGGGAGCTGGGGCTGGCCAGGACCAAGCTAGGGACCG	1080
PRIOR FILING DATE: 1997-09-19	Qy	1344	TACCCGCTCCCTTACACCGAGCGGTCTGATGAGGGGAGCGGCTGTGCGGCTGGT	1403
PRIOR FILING DATE: 1997-09-19	Db	1081	TACCCGCTCCCTTACACCGAGCGGTCTGATGAGGGGAGCGGCTGTGCGGCTGGT	1140
PRIOR FILING DATE: 1997-09-19	Qy	1404	GCCCATGGGAATACCCCGCACCTCTATGCGGACCAACCCGCTTCGAGGGGTACACCTGCC	1463
PRIOR FILING DATE: 1997-09-19	Db	1141	GCCCATGGGAATACCCCGCACCTCTATGCGGACCAACCCGCTTCGAGGGGTACACCTGCC	1200
PRIOR FILING DATE: 1997-09-19	Qy	1464	CAAGGGACGAGAGTCTTCCCTCTCTGCTGCTGATGAGGACCAAGCTTCAGGAGCA	1523
PRIOR FILING DATE: 1997-09-19	Db	1201	CAAGGGACGAGAGTCTTCCCTCTCTGCTGCTGATGAGGACCAAGCTTCAGGAGCA	1260
PRIOR FILING DATE: 1997-09-19	Qy	1524	GCACCCAGAGAGTTCACACCGAGCGGTTCCTGATGAGGACCAAGCTTCAGGAGCA	1583
PRIOR FILING DATE: 1997-09-19	Db	1261	GCACCCAGAGAGTTCACACCGAGCGGTTCCTGATGAGGACCAAGCTTCAGGAGCA	1320
PRIOR FILING DATE: 1997-09-19	Qy	1584	TGAGGGCTTCTGCTGCTTCTCTAGGAAAGCGTGTCTGCTTGGAGAGGCGCTGGCAAA	1643
PRIOR FILING DATE: 1997-09-19	Db	1321	TGAGGGCTTCTGCTGCTTCTCTAGGAAAGCGTGTCTGCTTGGAGAGGCGCTGGCAAA	1380
PRIOR FILING DATE: 1997-09-19	Qy	1644	AGCGGAGCTCTTCTCTCTTCTTACACCATCTTCTGCTGCTGAGGAGGCGCTGGT	1703
PRIOR FILING DATE: 1997-09-19	Db	1381	AGCGGAGCTCTTCTCTCTTCTTACACCATCTTCTGCTGCTGAGGAGGCGCTGGT	1440
PRIOR FILING DATE: 1997-09-19	Qy	1704	CCGCGCGAGACCTCTGAGGCTCAAGCCACCGCTCAGTGGCTTTCACATTTCCCGCAGC	1763
PRIOR FILING DATE: 1997-09-19	Db	1441	CCGCGCGAGACCTCTGAGGCTCAAGCCACCGCTCAGTGGCTTTCACATTTCCCGCAGC	1500

Remaining Prior Application data removed - See File Wrapper or PALM.

Query Match 89.0%; Score 2621.4; DB 13; Length 2623;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2623; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy	264	GCGGAGCGCTGGAGAGGAGGAGGAGCGACCTGCGGATGGAGCGGACCGCACCTG	323
Db	1	GCGGAGCGCTGGAGAGGAGGAGGAGCGACCTGCGGATGGAGCGGACCGCACCTG	60
Qy	324	GCGGCTGCTGCTGCGCTGCGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	383
Db	61	GCGGCTGCTGCTGCGCTGCGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	120
Qy	384	GCGGCGAGGCGACCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	443
Db	121	GCGGCGAGGCGACCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	180
Qy	444	GCTACGCGCGGCGGCGGCTGATTCAGGGCTCATGCGGCTGAGTAAAGTACGAGCGGT	503
Db	181	GCTACGCGCGGCGGCGGCTGATTCAGGGCTCATGCGGCTGAGTAAAGTACGAGCGGT	240
Qy	504	GTTCCACATCTACTGGGACCTCGGCGGCTGTGGTGTCTGCTGCTGCTGCTGCTGCTG	563
Db	241	GTTCCACATCTACTGGGACCTCGGCGGCTGTGGTGTCTGCTGCTGCTGCTGCTGCTG	300
Qy	564	GCGGAGGCGCTTGGAGGTACGCTGAGGAGTTCAGGGCGGCGGCGGCGGCGGCGGCG	623
Db	301	GCGGAGGCGCTTGGAGGTACGCTGAGGAGTTCAGGGCGGCGGCGGCGGCGGCGGCG	360
Qy	624	GGAAGGAGCTTTGATGGCCATGGGTTTCTTCTCCAAAGCGGAGCGGTGGAGGACGT	683









Db	2041	TGCTATGGCCCCAACTCATGCTCCCTCTCTTGGCTACACCACTCTCCCGAGCTGTGACC	2100
Qy	2364	ACCGATGTCACACACACCCCAACCACTTGTCCACACAGCTACCCAGTACGACATCGTCC	2423
Db	2101	ACCGATGTCACACACACCCCAACCACTTGTTCACACAGCTACCCAGTACAAACATCGTCC	2160
Qy	2424	TGGCTCCCCAGAGTATCTTCCCACTGAGACACGCGCGCCCCACAGAGGCACAGTCCCCAG	2483
Db	2161	TGGCTCCCCAGAGTATCTTCCCACTGAGACACGCGCGCCCCACAGAGGCACAGTCCCCAG	2220
Qy	2484	CCACTCTGCAACTGCGAGCGCCTCAGTACACCCCTTTTAAAGCACCCCTGATTCTACCAAAATG	2543
Db	2221	CCACTCTGCAACTGCGAGCGCCTCAGTACACCCCTTTTAAAGCACCCCTGATTCTACCAAAATG	2280
Qy	2544	CAAAACATCTGGGCTCTGCGATTATGCACAGAGACTTTGGACATACGAGGACCTCAGAC	2603
Db	2281	CAAAACATCTGGGCTCTGCGATTATGCACAGAGACTTTGGACATACGAGGACCTCAGAC	2340
Qy	2604	CGGAGGAACACTCTGCCCAAACCCCAACACGTCCTTATGTAAACACGTTGGAAGCGGCCCT	2663
Db	2341	CGGAGGAACACTCTGCCCAAACCCCAACACGTCCTTATGTAAACACGTTGGAAGCGGCCCT	2400
Qy	2664	GCTGCCCTCCACACACATACACACTCTGATCTACAGCCCTGTTCTCGGCGTCAGAG	2723
Db	2401	GCTGCCCTCCACACACATACACACTCTGATCTACAGCCCTGTTCTCGGCGTCAGAG	2460
Qy	2724	TCCCCACTAGACCCAGTGGAGGGGTTAGAGACCAAGTAGGGGCCAGTTTCCAAATCACC	2783
Db	2461	TCCCCACTAGACCCAGTGGAGGGGTTAGAGACCAAGTAGGGGCCAGTTTCCAAATCACC	2520
Qy	2784	CTGTCCAGGAGTGAGCCGGATCTCAACGTTCTTGTGACTTAAAGGGTCGGCTTGGGAATT	2843
Db	2521	CTGTCCAGGAGTGAGCCGGATCTCAACGTTCTTGTGACTTAAAGGGTCGGCTTGGGAATT	2580
Qy	2844	AAAGTTGTTTCTGGCCTTTAGCCCTAAAAAATAAAAAAAAAA	2886
Db	2581	AAAGTTGTTTCTGGCCTTTAGCCTAAAAAATAAAAAAAAAA	2623

Search completed: September 18, 2004, 01:40:01  
Job time : 1312.36 secs

[illegible]

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 17, 2004, 07:12:45 ; Search time 7174.05 Seconds  
(without alignments)  
12254.481 Million cell updates/sec

Title: US-10-669-693-1

Perfect score: 2944

Sequence: I ttttttgtttgttactc.....aaaaaaaaaaaaaaaaaaaaa 2944

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

```

1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estom:*
16: em_estfun:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1045.2	35.5	2621	11 AK004699	AK004699 Mus muscu
2	1045	35.5	2292	11 AK087069	AK087069 Mus muscu
3	884.8	30.1	900	13 BU552396	BU552396 AGENCOURT
4	871	29.6	884	13 BQ928732	BQ928732 AGENCOURT

5	867	29.4	899	13	BQ955833
6	865.2	29.4	945	13	BQ920829
7	861.6	29.3	924	13	BU849902
8	848.4	28.8	903	13	BQ900302
9	833	28.3	915	13	BU552425
10	811	27.5	877	12	BM044271
11	807.8	27.4	939	13	BQ958112
12	801	27.2	857	12	BM047426
13	783.6	26.6	804	12	BM044172
14	764.8	26.0	877	13	BU543186
15	716.2	24.3	969	13	BU543186
16	715	24.3	726	12	BM771940
17	707.8	24.0	838	13	BQ957100
18	702.8	23.9	729	13	BQ108766
19	687.2	23.3	2742	11	AK054324
20	673.6	22.9	924	14	CD049583
21	656.4	22.3	659	12	BQ001642
22	648	22.0	704	12	BM048532
23	646	21.9	646	12	BM739256
24	645	21.9	645	12	BM771818
25	621.6	21.1	879	12	BM046427
26	612	20.8	612	12	BM771424
27	612	20.8	626	12	BM766879
28	611.2	20.8	723	12	BI906340
29	602	20.4	602	12	BM756232
30	601.4	20.4	606	12	BM751841
31	599	20.3	599	12	BM786289
32	585.2	19.9	775	13	BU941997
33	574	19.5	574	12	BG123000
34	573.8	19.5	577	12	BM752062
35	570.4	19.4	572	12	BM850926
36	567	19.3	567	12	BM743865
37	566	19.2	578	12	BM747387
38	560.2	19.0	589	12	BM746882
39	554.4	18.8	648	12	BI016305
40	554	18.8	554	12	BI016305
41	553	18.8	553	12	BM771002
42	549.8	18.7	553	12	BM771294
43	548	18.6	548	13	BX646408
44	544.4	18.5	571	10	BE148597
45	541	18.4	541	12	BM752045

#### ALIGNMENTS

RESULT 1  
AK004699

LOCUS  
DEFINITION

AK004699 2621 bp mRNA linear HTC 20-SEP-2003  
Mus musculus adult male lung cDNA, RIKEN full-length enriched library, clone:1200011C15 product:similar to CYTOCHROME P450 2S1 [Homo sapiens], full insert sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

BQ955833 AGENCOURT  
BQ920829 AGENCOURT  
BU849902 AGENCOURT  
BQ900302 AGENCOURT  
BU552425 AGENCOURT  
BM044271 603621541  
BQ958112 AGENCOURT  
BM047426 603628712  
BM044172 603621408  
BU543186 AGENCOURT  
BX422464 BX422464  
BM771940 K-EST0055  
BQ957100 AGENCOURT  
BQ108766 imageqc 9  
AK054324 Mus muscu  
CD049583 AGENCOURT  
BQ001642 UI-H-DH1-  
BM048532 603625051  
BM739256 K-EST0008  
BM771818 K-EST0055  
BM046427 603626457  
BM771424 K-EST0055  
BM766879 K-EST0048  
BI906340 603063307  
BM756232 K-EST0034  
BM751841 K-EST0028  
BM786289 K-EST0064  
BU941997 AGENCOURT  
BG123000 603348696  
BM752062 K-EST0028  
BM850926 K-EST0131  
BM743865 K-EST0017  
BM747387 K-EST0021  
BM746882 K-EST0021  
BI016305 MR4-ET013  
BM771002 K-EST0054  
BM771294 K-EST0055  
BM751056 K-EST0026  
BX646408 DXFzp781K  
BE148597 MR0-HT024  
BM752045 K-EST0028

REFERENCE AUTHORS	3	Shibata, K., Itoh, M., Aizawa, K., Nagao, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Katsunari, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.
TITLE		RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multiplex capillary sequencer
JOURNAL MEDLINE PUBMED	20530913 11076861	Genome Res. 10 (11), 1757-1771 (2000)
REFERENCE AUTHORS	4	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
TITLE		Functional annotation of a full-length mouse cDNA collection
JOURNAL REFERENCE AUTHORS	5	Nature 409, 685-690 (2001)
TITLE		The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
JOURNAL REFERENCE AUTHORS	6	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs Nature 420, 563-573 (2002)
TITLE		6 (bases 1 to 2621)
JOURNAL REFERENCE AUTHORS		Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Arai, A., Aono, H., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Saigo, H., Saigo, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, H., Toyai, T., Yamamura, T., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
TITLE		Direct Submission
JOURNAL		Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
COMMENT		On Dec 10, 2002 this sequence version replaced gi:12836062. Please visit our web site (http://genome.gsc.riken.go.jp/) for further details. cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer 15' GAGAGAGAGCGCCGCACTCGAGTTTCTTTTCTTTTCTTNN 3', cDNA was prepared by using triose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence 5' GAGAGAGAGAGATCCAGAGCTCAATTAATTAATTAACCCCGCC 3'. cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: SstI; 3' end: XhoI. Host: SOLR. Location/Qualifiers 1. 2621 /organism="Mus musculus" /mol_type="mRNA" /strain="C57BL/6J" /db_xref="PANTOM_DB:1200011C15" /db_xref="WGI:1907239" /clone="taxon:10090" /clone="1200011C15" /sex="male" /tissue type="lung" /clone_lib="RIKEN full-length enriched mouse cDNA library" /dev_stage="adult"
FEATURES Source		1. 2621 /organism="Mus musculus" /mol_type="mRNA" /strain="C57BL/6J" /db_xref="PANTOM_DB:1200011C15" /db_xref="WGI:1907239" /clone="taxon:10090" /clone="1200011C15" /sex="male" /tissue type="lung" /clone_lib="RIKEN full-length enriched mouse cDNA library" /dev_stage="adult"
CDS		29. .1534 /note="unnamed protein product; putative similar to CYTOCHROME P450 2S1 [Homo sapiens] (SPT1Q9B266, evidence: FASTY, 78.4%ID, 99%length, match=1500)." /codon_start=1 /protein_id="BAB23484.1" /db_xref="GI:12836063" /translation="MEAASTWAILLALLLLLSLFRTPARGYIPPGTPTPLLG NLQLRFGALYSGLRLSKYGPVTYVLPWRRVVLVGHDAVREALGQAEFSGR GTIATLDFTDGHGVFFANGERWQLKFTLLALRDLGMKREGEELIQAQVSLVEA FOKEGRFPNPSMLLAQATSNVCSLVFGIRLPDDKFEQFQVIAQASTLLAGISPMW QAYEMFSLIQLPPLPHTQLQHLGTLAAFTIQVQKHQFQFSGPARDVDFLLX MAOBKODPTEFTENKMLMTVYLLFAGTWTIGATVYALLLLRLYPOVQORREELI OELGGRAPSLSDRVRLPYTDVLAHEAQRLLALVPMGPHPTITRTTCFRGTLPKTE VFPLIGSLHDHPAVFQNPGRFDELDGLRKEAFLPYSLGRKVLCEGLARAE LWLFPTSLQAFSLPTCPGDDLSLKPAISLGFNIIPDPFQLRVMTPTGQSR" 2600. .2605 /note="putative" 2621 /note="putative"
polyA_signal		polyA_site
ORIGIN		Query Match 35.5%; Score 1045.2; DB 11; Length 2621; Best Local Similarity 80.9%; Pred. No. 1.7e-122; Matches 1231; Conservative 0; Mismatches 288; Indels 3; Gaps 1; QY 283 AGAAGAGCCGACCTGCCGAGATGAGGACCCGACCTGGCGCTGCTGCTGGC---G 339 DB 8 AGACCCAGCCGACCTGCCGAGATGAGGACCCGACCTGGCGCTGCTGCTGGCCTG 67 QY 340 CTGGCGCTGCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 399 DB 68 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 127 QY 400 CCCCCCGGGCCCAACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 459 DB 128 CCCCCCGGGCCCAACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 187 QY 460 CTGTATTACGGGCTCATCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 519 DB 188 CTGTATTACGGGCTCATCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 247 QY 520 GGACCTGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 579 DB 248 GGCCCTGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 307 QY 580 GGTTCAGGCTGAGGAGTTTACGGCCCGGGGACCGTAGCGATGCTGGAAGGAGCTTTTGAT 639 DB 308 GGTTCAGGCTGAGGAGTTTACGGCCCGGGGACCGTAGCGATGCTGGAAGGAGCTTTTGAT 367 QY 640 GGCCATGGGGTTTCTTCTTCCACGGGCGGCTGCGAGGCGAGCTGAGGAGAGTTTACCATG 699 DB 368 GGTTCAGGAGTTTCTTCTTCCATGGGAGCGGTGGGAAACAGCTGAGGAAATCAACCTG 427 QY 700 CTTCGCTCTGCGGACCTGGGCGATGGGAAACGAGAGGCGAGGAGCTGATCCAGGCGGAG 759 DB 428 CTTCGCTCTGCGGACCTGGGCGATGGGCGATGGGCGAGGAGGAGGAGCTGATCCAGGCGGAG 487 QY 760 GCCCGGTGCTGCTGGAGACATTCAGGGGACAGAGGAGCGCCATTCGATCCCTCCCTG 819 DB 488 GTGCAGAGTCTGGTGGAGGCTTCCAGAGAGAGAGGAGCGTCCATTCACCCCTTCATG 547 QY 820 CTGCTGCCCGCAGGCGACCTCCAGAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 879 DB 548 CTGCTGCCCGCAGGCGACCTCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 607 QY 880 TATCAGGATAGGAGTTTCCAGGCGGCTGGTCCGGGCGAGCTGGTGTACCTGCTGGGAGTC 939 DB 608 TATCAGGATAGGAGTTTCCAGGCGGCTGGTCCAGGCGAGGAGGAGTGTACCTGTTGGGATC 667 QY 940 AGCTCCCGGCGGCTCAGACTACAGATGTTCTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 999



Db 668 AGCTCTCCATGGGGCCAGCCCTACGAGATGTTCTCTGGCTACTGACGCCCTGCCAGGC 727  
QY 1000 CCCACAGAGCTCTCCACAGCTCAGACCTTGGCTGCTTCAAGTCCGGAGGTG 1059  
Db 728 CCCACACAGCTCCAGCACCTTTGGGCGCCCTGGCTGCTTCACTATCCAGCAGGTA 787  
QY 1060 CAGCAGCACAGGGGAACTCGATGCTTTCGGGGCCCGCACGTGACCTTGTGATGCCCTTC 1119  
Db 788 CAGAAACACAGGAGGACCTTCCAAACCTCAGTCTTGACAGTGTGCTGTTGACGCCCTTC 847  
QY 1120 CTGCTGAAGTGGACAGGAGGAAACAAACCCAGGACAGAAATTCACCAACAAGAACATG 1179  
Db 848 CTGCTGAAGTGGACAGGAGGAAACAAAGCCAGGTACAGAAATTCACCGAGAAGACTTG 907  
QY 1180 CTGATGACAGTCATTTATTTGCTGTTTCTGGGAGGATGAGGTGACGACCCACCGTGGC 1239  
Db 908 CTGATGAGGTGACATACCTGCTGTTTCTGGGACCATGACCATCGGTGGCCACCATCCG 967  
QY 1240 TATACCTCTGCTCTGTGATGAATACCTCATGTCCAAAGTGGGTACGTGAGGAGCTG 1299  
Db 968 TATGCCCTCTGCTCTGTGATGATACCTCAAGTCCAGCAGCGCTCGGGAGGAGCTC 1027  
QY 1300 AATCGGAGCTGGGGTGGCCAGGACCAAGCCTAGGGACCGTACCCGCTCCCTTAC 1359  
Db 1028 ATACGAGGCTGGGTCTCGTGGAGGGCTCCAACTCTCAGCGATCGAGTTCGCTCCCTTAC 1087  
QY 1360 ACCGACCGCTTCTGATGAGGCGCAGCGCTGCTGGCGTGGTGGCCATGGGATACCC 1419  
Db 1088 ACGATGCGCTTTACAGAGGACACGCGCTCTCTGGCATCTGGTACCATGGGATGCC 1147  
QY 1420 CGACCTCATGCGGACCAACCGCTTCCGAGGGTACACCTGCCCCAGGCGACGAGGTC 1479  
Db 1148 CACACCATCAGGAGGACCACTTGTCTCCGAGGGTACACTCTGCCAGGSCATGAGTC 1207  
QY 1480 TTCCCTCTGCTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1539  
Db 1208 TTCCCTCTGATGGCTTCACTACTGATGATGATGATGATGATGATGATGATGATGATG 1267  
QY 1540 AACCCAGACGCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1599  
Db 1268 CATCCAGGCGCTTCTGACGAGGATGATGATGATGATGATGATGATGATGATGATG 1327  
QY 1600 TTCTCTTGGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1659  
Db 1328 TACTCTTGGTAAAGGAGTGTCTGCTGAGAGGCTGCTGCTGAGGAGTGTGCTGCTGCT 1387  
QY 1660 TTCTTACCACCATCTACAGCTTCTGCTGAGAGGCTGCTGCTGAGGAGGCTGCTGCTG 1719  
Db 1388 TTCTTCACTTCCATCTTGAAGCTTCTGCTGAGAGGCTGCTGCTGAGAGGCTGCTGCTG 1447  
QY 1720 AGCTTCAAGCCACCTCAGTGGCTTTTCAATTTCCATTTCCCGGCTTCCAGCTGGAAGTC 1779  
Db 1448 AGCTTCAAGCCACCTCAGTGGCTTTTCAATTTCCATTTCCCGGCTTCCAGCTGGAAGTC 1507  
QY 1780 GTTCCCATGACTTCACTCCA 1801  
Db 1508 TGGCCCACTGGGACAGGCTCA 1529

RESULT 2  
AK087069  
LOCUS AK087069 2292 bp mRNA linear HTC 20-SEP-2003  
DEFINITION Mus musculus 0 day neonate lung cDNA, RIKEN full-length enriched library, clone:E030025C01 product:similar to CYTOCHROME P450 2S1 (Homo sapiens), full insert sequence.  
ACCESSION AK087069  
VERSION AK087069.1 GI:26352314  
KEYWORDS HTC; CAP trapper.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1

## AUTHORS

TITLE  
JOURNAL  
MEDLINE  
PUBMED

## AUTHORS

TITLE  
JOURNAL  
MEDLINE  
PUBMED

## AUTHORS

TITLE  
JOURNAL  
MEDLINE  
PUBMED

## AUTHORS

TITLE  
JOURNAL  
MEDLINE  
PUBMED

## AUTHORS

TITLE  
JOURNAL  
MEDLINE  
PUBMED

## AUTHORS

TITLE  
JOURNAL  
MEDLINE  
PUBMED

## AUTHORS

TITLE  
JOURNAL  
MEDLINE  
PUBMED

## AUTHORS

TITLE  
JOURNAL  
MEDLINE  
PUBMED

## AUTHORS

TITLE  
JOURNAL  
MEDLINE  
PUBMED

## AUTHORS

TITLE  
JOURNAL  
MEDLINE  
PUBMED

## AUTHORS

TITLE  
JOURNAL  
MEDLINE  
PUBMED

## AUTHORS

TITLE  
JOURNAL  
MEDLINE  
PUBMED

## AUTHORS

TITLE  
JOURNAL  
MEDLINE  
PUBMED

## AUTHORS

TITLE  
JOURNAL  
MEDLINE  
PUBMED

## AUTHORS

TITLE  
JOURNAL  
MEDLINE  
PUBMED

## AUTHORS

TITLE  
JOURNAL  
MEDLINE  
PUBMED

## AUTHORS

TITLE  
JOURNAL  
MEDLINE  
PUBMED

## AUTHORS

TITLE  
JOURNAL  
MEDLINE  
PUBMED

## AUTHORS

TITLE  
JOURNAL  
MEDLINE  
PUBMED

## AUTHORS

TITLE  
JOURNAL  
MEDLINE  
PUBMED

## AUTHORS

TITLE  
JOURNAL  
MEDLINE  
PUBMED

## AUTHORS

TITLE  
JOURNAL  
MEDLINE  
PUBMED

## AUTHORS

TITLE  
JOURNAL  
MEDLINE  
PUBMED

## AUTHORS

TITLE  
JOURNAL  
MEDLINE  
PUBMED

## AUTHORS

TITLE  
JOURNAL  
MEDLINE  
PUBMED

## AUTHORS

TITLE  
JOURNAL  
MEDLINE  
PUBMED

## AUTHORS

TITLE  
JOURNAL  
MEDLINE  
PUBMED

## AUTHORS

TITLE  
JOURNAL  
MEDLINE  
PUBMED

## AUTHORS

TITLE  
JOURNAL  
MEDLINE  
PUBMED

## AUTHORS

TITLE  
JOURNAL  
MEDLINE  
PUBMED

## AUTHORS

TITLE  
JOURNAL  
MEDLINE  
PUBMED

## AUTHORS

TITLE  
JOURNAL  
MEDLINE  
PUBMED

## AUTHORS

TITLE  
JOURNAL  
MEDLINE  
PUBMED

## AUTHORS

TITLE  
JOURNAL  
MEDLINE  
PUBMED

## AUTHORS

TITLE  
JOURNAL  
MEDLINE  
PUBMED

## AUTHORS

TITLE  
JOURNAL  
MEDLINE  
PUBMED

## AUTHORS

TITLE  
JOURNAL  
MEDLINE  
PUBMED

## AUTHORS

TITLE  
JOURNAL  
MEDLINE  
PUBMED

## AUTHORS

TITLE  
JOURNAL  
MEDLINE  
PUBMED

## AUTHORS

TITLE  
JOURNAL  
MEDLINE  
PUBMED

## AUTHORS

TITLE  
JOURNAL  
MEDLINE  
PUBMED

## AUTHORS

TITLE  
JOURNAL  
MEDLINE  
PUBMED

## AUTHORS

TITLE  
JOURNAL  
MEDLINE  
PUBMED

## AUTHORS

TITLE  
JOURNAL  
MEDLINE  
PUBMED

## AUTHORS

TITLE  
JOURNAL  
MEDLINE  
PUBMED

## AUTHORS

TITLE  
JOURNAL  
MEDLINE  
PUBMED

## AUTHORS

TITLE  
JOURNAL  
MEDLINE  
PUBMED

Carninci, P. and Hayashizaki, Y.  
High-efficiency full-length cDNA cloning  
Meth. Enzymol. 303, 19-44 (1999)  
99279253  
10349636

2  
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,  
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new genes  
Genome Res. 10 (10), 1617-1630 (2000)  
20499374  
11042159

3  
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,  
Konno, H., Akiyama, J., Nishi, K., Kitsuunai, T., Tashiro, H., Itoh, M.,  
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,  
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,  
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,  
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,  
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
RIKEN integrated sequence analysis (RISA) system--384-format  
sequencing pipeline with 384 multicapillary sequencer  
Genome Res. 10 (11), 1757-1771 (2000)  
20530913  
11078861

4  
The RIKEN Genome Exploration Research Group Phase II Team and the  
FANTOM Consortium.  
Functional annotation of a full-length mouse cDNA collection  
Nature 409, 685-690 (2001)

5  
The FANTOM Consortium and the RIKEN Genome Exploration Research  
Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation  
of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

6 (bases 1 to 2292)  
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,  
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,  
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,  
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,  
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,  
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,  
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,  
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,  
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,  
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,  
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,  
Muramatsu, M. and Hayashizaki, Y.  
Direct Submission  
Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of  
Physical and Chemical Research (RIKEN), Laboratory for Genome  
Exploration and Research Group, RIKEN Genomic Sciences Center (GSC),  
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,  
Kanagawa 230-0045, Japan (E-mail: genome-res@sc.riken.go.jp,  
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,  
Fax: 81-45-503-9216)

cDNA library was prepared and sequenced in Mouse Genome  
Encyclopedia Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
Division of Experimental Animal Research in Riken contributed to  
prepare mouse tissues.  
Please visit our web site for further details.  
URL: http://genome.gsc.riken.go.jp/  
URL: http://fantom.gsc.riken.go.jp/  
Location/Qualifiers  
1. .2292  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="FANTOM\_DB:E030025C01"  
/db\_xref="MGI:2424609"

FEATURES  
source



AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg. Ph. D.

Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)  
Tissue Procurement: DCTD/DTP  
CDNA Library Preparation: Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLCM2774 row: e column: 06  
High quality sequence stop: 800.

```

FEATURES
source
      1. .900
      Location/Qualifiers
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IWAGE:6576174"
            /tissue_type="carcinoma, cell line"
            /lab_host="DH10B (phage-resistant)"
            /clone_lib="NIH_MGC_40"
            /notes="Organ: prostate; Vector: pOTB7; Site 1: XhoI;
            Site 2: EcoRI; cDNA made by oligo-dT priming.
            Directionally cloned into EcoRI/XhoI sites using the
            following 5' adaptor: GGCACGAG(G). Library constructed by
            Ling Hong in the laboratory of Gerald M. Rubin (University
            of California, Berkeley) using ZAP-cDNA synthesis kit
            (Stratagene) and Superscript II RT (Life Technologies).
            Note: this is a NIH MGC Library."

```

ORIGIN

Note: This is a NIH\_MGC Library.

Query Match	30.1%	Score 884.8;	DB 13;	Length 900;
Best Local Similarity	99.7%;	Pred. No. 3.8e-102;		
Matches 997;	Conservative 0;	Mismatches 2;	Indels 1;	Gaps 1;

645	QY	TGGGGTTTTCTTCTCCAA	CGGGGAGCGGTGGAGC	CAGCTCAGGAAGTTTACCA	TGCTTGC	704	
	Db						
1	QY	TGGGGTTTTCTTCTCC	A	CGGGGAGCGGTGGAGC	AGCTCAGGAAGTTTACCA	TGCTTGC	59
	Db						
705	QY	TCTCGGGACCTTGGGCAT	TGGGAAGCGAGAAGCG	CAGGAGCTGATCAGCGGAGGCGCCG		764	
	Db						
60	QY	TCTCGGGACCTTGGGCAT	TGGGAAGCGAGAAGCG	CAGGAGCTGATCAGCGGAGGCGCCG		119	
	Db						
765	QY	GTGCTGTGTGGAGACATT	CCAGGGGACAGAAAGAC	CGCCCATTCGATCCCTCCCTGTGCT		824	
	Db						
120	QY	GTGCTGTGTGGAGACATT	CCAGGGGACAGAAAGAC	CGCCCATTCGATCCCTCCCTGTGCT		179	
	Db						
825	QY	GGCCACGGCCACCTC	CAACGTAGTCTGTCTCCCT	CTCTTTGGGCTCCGGCTTCCTCCATGA		884	
	Db						
180	QY	GGCCACGGCCACCTC	CAACGTAGTCTGTCTCCCT	CTCTTTGGGCTTCGGCTTCTCCATGA		239	
	Db						
885	QY	GGATAAGAGATTCCAG	CGCCGTGGTCCGGGAC	AGTGTGTGATACCTCTGTGGAGTCAGCTC		944	
	Db						
240	QY	GGATAAGAGATTCCAG	CGCCGTGGTCCGGGAC	AGTGTGTGATACCTCTGTGGAGTCAGCTC		299	
	Db						
945	QY	CCAGGGGGTTCAGACT	ACGAGATGTTCTTCC	TGGTTCCTGGGCCCTTCGCAGGCCCCCA		1004	
	Db						
300	QY	CCAGGGGGTTCAGACT	ACGAGATGTTCTTCC	TGGTTCCTGGGCCCTTCGCAGGCCCCCA		359	
	Db						
1005	QY	CAAGCAGCTCCTCCAC	CCAGCTCAGACACCTTGG	CTGGCTTCACAGTCCGGCAGGTGCAGCA		1064	
	Db						
360	QY	CAAGCAGCTCCTCCAC	CCAGCTCAGACACCTTGG	CTGGCTTCACAGTCCGGCAGGTGCAGCA		419	
	Db						
1065	QY	GCAACAGGGGAACCTT	GGGATGCTTGGGGCCCCG	ACGTGACCTTGTGCATGCCTTCCCTGCT		1124	
	Db						
420	QY	GCAACAGGGGAACCTT	GGGATGCTTGGGGCCCCG	ACGTGACCTTGTGCATGCCTTCCCTGCT		479	
	Db						
1125	QY	GAGATGACACAGGAGGA	CAAAACCCAGGCA	CAGAAATTCACCAACAGAA	CATGCTGAT	1184	
	Db						
480	QY	GAGATGACACAGGAGGA	CAAAACCCAGGCA	CAGAAATTCACCAACAGAA	CATGCTGAT	539	
	Db						

Qy	1185	GACAGTCATTTATTTGCTGTTTGTCTGGACGATGACGGTCAGCACACCGTCGGCTATAC	1244
Db	540	GACAGTCATTTATTTGCTGTTTGTCTGGACGATGACGGTCAGCACACCGTCGGCTATAC	599
Qy	1245	CCTCCTGCTCCTGATGAAATACCCCTCATGTCCAAAGTGGGTACGTGAGGAGCTGAATCG	1304
Db	600	CCTCCTGCTCCTGATGAAATACCCCTCATGTCCAAAGTGGGTACGTGAGGAGCTGAATCG	659
Qy	1305	GGAGCTGGGGGCTGGCCAGGACCAACAGCCTAGGGGACCGTACCCGCTCCCTTACACCGA	1364
Db	660	GGAGCTGGGGGCTGGCCAGGACCAACAGCCTAGGGGACCGTACCCGCTCCCTTACACCGA	719
Qy	1365	CGCGGTTCTGTCATGAGGCGCAGCGGCTGCTGGCGCTGGTGCCCATGGGAATACCCCGCAC	1424
Db	720	CGCGGTTCTGTCATGAGGCGCAGCGGCTGCTGGCGCTGGTGCCCATGGGAATACCCCGCAC	779
Qy	1425	CCTCATGCGGACCAACCGCTTCGAGGGTACACCTTCGCCCGCAGGGACCGAGGCTCTTCCC	1484
Db	780	CCTCATGCGGACCAACCGCTTCGAGGGTACACCTTCGCCCGCAGGGACCGAGGCTCTTCCC	839
Qy	1485	CCTCCTTGGCTCCATCCTGTCATGACCCCAACATCTTCAAGCACCCAGAGAGTTCAACCC	1544
Db	840	CCTCCTTGGCTCCATCCTGTCATGACCCCAACATCTTTCAGCACCCAGAGAGTTCAACCC	899

RESULT 4	BQ928732	LOCUS	BQ928732	884 bp	mRNA	linear	EST 20-AUG-2002
DEFINITION	AGENCOURT_10037033	NIH MGC 40	Homo sapiens	cDNA clone IMAGE:6480553	5', mRNA sequence.		

ACCESSION	BQ928732	
VERSION	BQ928732.1	GI:22343763
KEYWORDS	EST.	
SOURCE	Homo sapiens	(human)
ORGANISM	Homo sapiens	

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 884)  
NIH-MGC <http://mgc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.

Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)  
Tissue Procurement: DCTD/DTF  
cDNA Library Preparation: Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLCM2660 row: m column: 06  
High quality sequence stop: 792.

FEATURES  
source

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6480653"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_40"
/notes="Organ: prostate; Vector: pOTB7; Site 1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH MGC Library."

```

## ORIGIN

Query Match 29.6%; Score 871; DB 13; Length 884;  
Best Local Similarity 99.8%; Pred. No. 2.1e-100;

Matches	882;	Conservative	0;	Mismatches	1;	Indels	1;	Gaps	1;
Qy	1323	GGCACCAGCCTAGGGACCGTACCGGCTCCCTTACACCGACGGGTTCTGCAATGAGC	1382						
Db	1	GGCACCAGCCTAGGGACCGTACCGGCTCCCTTACACCGACGGGTTCTGCAATGAGC	60						
Qy	1383	GCAGGGGTGCTGGCGCTGGTGCCTATGGGAATACCCGCGACCTCTATGCGGACACCG	1442						
Db	61	GCAGGGGTGCTGGCGCTGGTGCCTATGGGAATACCCGCGACCTCTATGCGGACACCG	120						
Qy	1443	CTTCGAGGGTACACCTGCGCCCGACGAGGCTCTTCCCTCTTGGCTTCCATCT	1502						
Db	121	CTTCGAGGGTACACCTGCGCCCGACGAGGCTCTTCCCTCTTGGCTTCCATCT	180						
Qy	1503	GCATGACCCCAACATCTTCAAGCACCCAGAGAGTTCAACCCAGACCGTTTCTTGATGC	1562						
Db	181	GCATGACCCCAACATCTTCAAGCACCCAGAGAGTTCAACCCAGACCGTTTCTTGATGC	240						
Qy	1563	AGATGGACGTTTCAGGAAGCATGAGCGTCTCTGCGCTTCTCTTAGGGAAGCGTGTG	1622						
Db	241	AGATGGACGTTTCAGGAAGCATGAGCGTCTCTGCGCTTCTCTTAGGGAAGCGTGTG	300						
Qy	1623	CCTTGGAGGGCCTGGCAAGACCGAGCTCTTCTCTTCTTTCACCAACATCTTACAAGC	1682						
Db	301	CCTTGGAGGGCCTGGCAAGACCGAGCTCTTCTCTTCTTTCACCAACATCTTACAAGC	360						
Qy	1683	CTTCTCCCTGGAGACCGGTGCCCGGACACCTGTAGCTTCAAGCCACCGTCAGTGG	1742						
Db	361	CTTCTCCCTGGAGACCGGTGCCCGGACACCTGTAGCTTCAAGCCACCGTCAGTGG	420						
Qy	1743	CCTTTTCAACATCTCCCGACGCTTCCAGTGAAGTCCGTCCCTCACTGACCTTCACTCCAC	1802						
Db	421	CCTTTTCAACATCTCCCGACGCTTCCAGTGAAGTCCGTCCCTCACTGACCTTCACTCCAC	480						
Qy	1803	CACGACACACGATGAAGGAAGCACTTGGAGTGGTGGTGCACGAGCGGTGCTCC	1862						
Db	481	CACGACACACGATGAAGGAAGCACTTGGAGTGGTGGTGCACGAGCGGTGCTCC	540						
Qy	1863	AGCCTCAACAGTGGGATGGACAGGTTAATGTCTCCAGAGTGTACATGCGAGGACCA	1922						
Db	541	AGCCTCAACAGTGGGATGGACAGGTTAATGTCTCCAGAGTGTACATGCGAGGACCA	600						
Qy	1923	CATTTCACCGCTGACGTTGTTTCCGGAGTCTGTCCACGCGCCACACGCTCACTTGAC	1982						
Db	601	CATTTCACCGCTGACGTTGTTTCCGGAGTCTGTCCACGCGCCACACGCTCACTTGAC	560						
Qy	1983	TCATGCTGCTAAGATGCACACCGCACACCCATACACAACTACAAGGGCCACAAAGCAAC	2042						
Db	661	TCATGCTGCTAAGATGCACACCGCACACCCATACACAACTACAAGGGCCACAAAGCAAC	720						
Qy	2043	TGCTGGTTAGTTTCCACAGACATTAATATATGTCATCTGCAATCACAAGCACATAGCC	2102						
Db	721	TGCTGGTTAGTTTCCACAGACATTAATATATGTCATCTGCAATCACAAGCACATAGCC	780						
Qy	2103	AGGTAAACCAACCACTCCCTGATCTGACGCCACACCT-GGGAGTCTGGCTGTCACT	2161						
Db	781	AGGTAAACCAACCACTCCCTGATCTGACGCCACACCTGGGGAGTCTGGCTGTCACT	840						
Qy	2162	TCACAGCCACAGAAACGGCCACACATGTTTCAGCTCACAGC	2205						
Db	841	TCACAGCCACAGAAACGGCCACACATGTTTCAGCTCACAGC	884						

RESULT 5  
BQ895833  
LOCUS BQ895833 899 bp mRNA linear EST 16-AUG-2002  
DEFINITION AGENCOURT 8675453 NIH\_MGC\_40 Homo sapiens cDNA clone IMAGE:6380072  
5', mRNA sequence.  
ACCESSION BQ895833  
VERSION BQ895833.1 GI:22287847  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 899)  
NIH-MGC <http://mgi.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue procurement: DCTD/BTP  
cDNA Library Preparation: Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Plate: LUCM2568 row: b column: 09  
High quality sequence stop: 644.  
Location/Qualifiers  
1..899  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6380072"  
/tissue\_type="carcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH MGC 40"  
/note="Organ: prostate; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH\_MGC Library."

FEATURES  
source

ORIGIN

Query Match	29.4%;	Score 867;	DB 13;	Length 899;
Best Local Similarity	99.1%;	Pred. No. 6.6e-100;		
Matches 892;	Conservative	0;	Mismatches 6;	Indels 2;
Qy	1323	GGCACCAGCCTAGGGACCGTACCGGCTCCCTTACACCGACGGGTTCTGCAATGAGC	1382	
Db	1	GGCACCAGCCTAGGGACCGTACCGGCTCCCTTACACCGACGGGTTCTGCAATGAGC	59	
Qy	1383	GCAGGGGTGCTGGCGCTGGTGCCTATGGGAATACCCGCGACCTCTATGCGGACACCG	1442	
Db	60	GCAGGGGTGCTGGCGCTGGTGCCTATGGGAATACCCGCGACCTCTATGCGGACACCG	119	
Qy	1443	CTTCCGAGGGTACACCTGCCCCAGGCGACGGAGGTCTTCCCTCTCTTGGGTCCATCT	1502	
Db	120	CTTCCGAGGGTACACCTGCCCCAGGCGACGGAGGTCTTCCCTCTCTTGGGTCCATCT	179	
Qy	1503	GCATGACCCCAACATCTTCAAGCACCCAGAGTTCAACCCAGACCGTTTCTGATGC	1562	
Db	180	GCATGACCCCAACATCTTCAAGCACCCAGAGAGTTCAACCCAGACCGTTTCTGATGC	239	
Qy	1563	AGATGAGCGTTTCAGGAAGCATGAGCGTCTTCCCTCTTCTTAGGGAAGCGTGTCTG	1622	
Db	240	AGATGAGCGTTTCAGGAAGCATGAGCGTCTTCCCTCTTCTTAGGGAAGCGTGTCTG	299	
Qy	1623	CCTTGGAGGGCCTGGCAAGACCGGCTCTTCTCTTCTTTCACCAACATCTTACAAGC	1682	
Db	300	CCTTGGAGGGCCTGGCAAGACCGGAGCTCTTCTCTTCTTTCACCAACATCTTACAAGC	359	
Qy	1683	CTTCTCCCTGGAGAGCGCTGCCCGCGACACCTTGAGCTCAAGCCACCGTCAGTGG	1742	
Db	360	CTTCTCCCTGGAGAGCGCTGCCCGCGACACCTTGAGCTCAAGCCACCGTCAGTGG	419	
Qy	1743	CTTTTTCACATTTCCCGCCAGCCTTCCAGCTGCAAGTCCGTCCCACTGACCTTCACTCCAC	1802	
Db	420	CTTTTTCACATTTCCCGCCAGCCTTCCAGCTGCAAGTCCGTCCCACTGACCTTCACTCCAC	479	
Qy	1803	CACGACACACGATGAAGGAAGCACTTGGAGTGGTGGTGCACGAGCGGTGCTCC	1862	



DEFINITION AGENCOURT 10441031 NIH MGC 109 Homo sapiens cDNA clone  
IMAGE:6598606 5', mRNA sequence.  
ACCESSION BU849902.1 GI:24034865  
VERSION BU849902.1 GI:24034865  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 924)  
NIH-MGC http://mgc.nci.nih.gov/.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: ATCC  
cDNA Library Preparation: Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLCM2831 row: k column: 22  
High quality sequence stop: 701.  
Location/Qualifiers  
1. 924  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6598606"  
/tissue type="teratocarcinoma, cell line"  
/lab host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_109"  
/note="Organ: ovary; Vector: pOTB7; Site 1: EcoRI; Site 2:  
XhoI; cDNA made by oligo-dT priming. Directionally cloned  
into EcoRI/XhoI sites using the following 5' adaptor:  
GGCAGGAG(G). Library constructed by Ling Hong in the  
laboratory of Gerald M. Rubin (University of California,  
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and  
Superscript II RT (Life Technologies). Note: this is a  
NIH\_MGC Library."

ORIGIN  
Query Match 29.3%; Score 861.6; DB 13; Length 924;  
Best Local Similarity 98.6%; Pred No. 3.1e-99;  
Matches 900; Conservative 0; Mismatches 10; Indels 3; Gaps 3;  
QY 290 GCGACCTGCCGAGATGAGGCGACCGGCACCTGGCGCGTCTGCTGGCGTGGCGCTGC 349  
DB 1 GCGACCTGCCGAGATGAGGCGACCGGCACCTGGCGCGTCTGCTGGCGTGGCGCTGC 60  
QY 350 TCCTGCTGCTGACGCTGGCGTGTCCGGACCGAGGCCCGAGGCCACCTGCCCGCGGGC 409  
DB 61 TCCTGCTGCTGACGCTGGCGTGTCCGGACCGAGGCCCGAGGCCACCTGCCCGCGGGC 120  
QY 410 CCAGCCGCTACCACTGCTGGGAACCTCTCAGCTACGGCCCGGGCGGTGATTACG 469  
DB 121 CCAGCCGCTACCACTGCTGGGAACCTCTCAGCTACGGCCCGGGCGGTGATTACG 180  
QY 470 GGCTCATGCGGTGAGTAAGAGTACGACCGGTGTACCATCTACCTGGGACCTGGC 529  
DB 181 GGCTCATGCGGTGAGTAAGAGTACGACCGGTGTACCATCTACCTGGGACCTGGC 240  
QY 530 GGCCTGTGGTGTCTGCTGGGCGAGGCTGTCCGGAGAGCCCTGGAGGTGAGGTG 589  
DB 241 GGCCTGTGGTGTCTGCTGGGCGAGGCTGTCCGGAGAGCCCTGGAGGTGAGGTG 300  
QY 590 AGGAGTTACGCGCGCGGGAACCGTAGCATCTCGAAGCGGACTTTTGATGCCATGGGG 649  
DB 301 AGGAGTTACGCGCGCGGGAACCGTAGCATCTCGAAGCGGACTTTTGATGCCATGGGG 360  
QY 650 TTTTCTTCTCCACCGGGGAGCGGTGGAGGACGCTGAGGAAGTTTACCATGCTTGTCTGTC 709

DB 361 TTTTCTTCTCCACCGGGGAGCGGTGGAGGACGCTGAGGAAGTTTACCATGCTTGTCTGTC 420  
QY 710 GGGACCTGGGCGATGGGGAAGCGAGAAGCGGAGGAGCTGATCCAGCGGAGGCCCGGTGTC 769  
DB 421 GGGACCTGGGCGATGGGGAAGCGAGAAGCGGAGGAGCTGATCCAGCGGAGGCCCGGTGTC 480  
QY 770 TGGTGGAGACATTCCAGGGGACAGAAGGACGCCCATTCGATCCCTCCTGCTGTGGGCC 829  
DB 481 TGGTGGAGACATTCCAGGGGACAGAAGGACGCCCATTCGATCCCTCCTGCTGTGGGCC 540  
QY 830 AGGCACCTCCAAAGTAGTCTGCTCCCTCTCTTTGGGCTCGCTTCTCTATGAGGATA 889  
DB 541 AGGCACCTCCAAAGTAGTCTGCTCCCTCTCTTTGGGCTCGCTTCTCTATGAGGATA 600  
QY 890 AGGAGTTCCAGGCGGTGTCGGGCGAGCTGTGGTGTACCTCTGTGGAGTACGCTCCACGG 949  
DB 601 AGGAGTTCCAGGCGGTGTCGGGCGAGCTGTGGTGTACCTCTGTGGAGTACGCTCCACGG 660  
QY 950 GGGGTTCAGACCTACGAGATGTTCTCTGTTCTCTGCGGCCCTGCGAGGCCGCCCAAGC 1009  
DB 661 GGGGTTCAGACCTACGAGATGTTCTCTGTTCTCTGCGGCCCTGCGAGGCCGCCCAAGC 720  
QY 1010 AGCTCTCTCCACGCTACGACCTTGGCTGCTTTCAGAGTCTGGGAGTGGAGGAGCACC 1069  
DB 721 AGCTCTCTCCACGCTACGACCTTGGCTGCTTTCAGAGTCTGGGAGTGGAGGAGCACC 780  
QY 1070 AGGGG-AACTGTGATGCTTCGGGCCCGCGAGCTGTGATGCTTCTGATGCTT-CCTGCTGAA 1127  
DB 781 AGGGGAAACCTGTGATGCTTCGGGCCCGCGAGCTGTGATGCTTCTGATGCTTCTGCTGTA 840  
QY 1128 GATGCCAGGAGGAACAAACCCAGGCGACAGAA-TTCACCAACAAGAACATGCTGATGA 1186  
DB 841 GAGGCGACCGGAGGACAAACCCAGGCGACAGAA-TTCACCAACAAGAACATGCTGATGA 900  
QY 1187 CAGTCATTATTATT 1199  
DB 901 CAGTCATTATTATT 913  
RESULT 8  
BQ900302  
LOCUS AGENCOURT\_8675726 NIH\_MGC\_40 Homo sapiens cDNA clone IMAGE:6380105  
DEFINITION 5', mRNA sequence.  
ACCESSION BQ900302  
VERSION BQ900302.1 GI:22292316  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 903)  
NIH-MGC http://mgc.nci.nih.gov/.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: DCTD/DTF  
cDNA Library Preparation: Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLCM2568 row: c column: 18  
High quality sequence stop: 690.  
Location/Qualifiers  
1. 903  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6380105"  
/tissue\_type="carcinoma, cell line"



/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_40"  
/note="Organ: prostate; Vector: pOTB7; Site\_1: XhoI;  
Site\_2: EcoRI; cDNA made by oligo-dT priming.  
Directionally cloned into EcoRI/XhoI sites using the  
following 5' adaptor: GGCACGAG(G). Library constructed by  
Ling Hong in the laboratory of Gerald M. Rubin (University  
of California, Berkeley) using ZAP-cDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies).  
Note: this is a NIH\_MGC Library."

ORIGIN

Query Match 28.8%; Score 848.4; DB 13; Length 903;  
Best Local Similarity 98.1%; Pred. No. 1.4e-97;  
Matches 888; Conservative 0; Mismatches 14; Indels 3; Gaps 3;

QY 645 TGGGGTTTCTTCTCCAAACGGGAGCGGTGGAGGAGCTGAGGAAGTTTACCATGCTTGC 704  
DB |||||||  
QY 705 TCTCGGGAGCTGGGCATGGGAAGCGAGAGCGAGCTGATCCAGGCGGAGGCCG 764  
DB |||||||  
QY 59 TCTCGGGAGCTGGGCATGGGAAGCGAGAGCGAGCTGATCCAGGCGGAGGCCG 118  
DB |||||||  
QY 765 GTGTCTGTGGAGACATTCAGGGGAGCAGAGGACGCCATTTCATCCCTCCCTGCTGCT 824  
DB |||||||  
QY 119 GTGTCTGTGGAGACATTCAGGGGAGCAGAGGAGCGCCATTTCATCCCTCCCTGCTGCT 178  
DB |||||||  
QY 825 GGCCCGACCTCCAACTAGTCTGCTCCTCTCTTTGGCTCCGCTTCTCTATGA 884  
DB |||||||  
QY 179 GGCCCGACCTCCAACTAGTCTGCTCCTCTCTTTGGCTCCGCTTCTCTATGA 238  
DB |||||||  
QY 885 GGATAGGAGTTCAGGCGCTGTGCGGCGAGCTGTGTACCTGCTGGAGTCACTC 944  
DB |||||||  
QY 239 GGATAGGAGTTCAGGCGCTGTGCGGCGAGCTGTGTACCTGCTGGAGTCACTC 298  
DB |||||||  
QY 945 CCAGGGGGTTCAGACCTACGAGATGTTCTCTGCTGCGGCCCTGCGAGCCGCCA 1004  
DB |||||||  
QY 299 CCAGGGGGTTCAGACCTACGAGATGTTCTCTGCTGCGGCCCTGCGAGCCGCCA 358  
DB |||||||  
QY 1005 CAAGCAGCTCTCCACAGCTGAGCCTTGGTGTCTTCAAGTCCGGAGTGACGA 1064  
DB |||||||  
QY 359 CAAGCAGCTCTCCACAGCTGAGCCTTGGTGTCTTCAAGTCCGGAGTGACGA 418  
DB |||||||  
QY 1065 GCACACGGGAACTGATCTCTCGGCCCGCGACGCTGCTGTGATGCTCTGCT 1124  
DB |||||||  
QY 419 GCACACGGGAACTGATCTCTCGGCCCGCGACGCTGCTGTGATGCTCTGCT 478  
DB |||||||  
QY 1125 GAAGATGGCAAGGAGGAAACAAACCCAGGACAGAAATTCACCAACAGAAATGCTGAT 1184  
DB |||||||  
QY 479 GAAGATGGCAAGGAGGAAACAAACCCAGGACAGAAATTCACCAACAGAAATGCTGAT 538  
DB |||||||  
QY 1185 GACAGTCATTTATTTCTGTTCTGGAGAGATGAGGTGAGCACCAGCTCGGCTATAC 1244  
DB |||||||  
QY 539 GACAGTCATTTATTTCTGTTCTGGAGAGATGAGGTGAGCACCAGCTCGGCTATAC 598  
DB |||||||  
QY 1245 CTTCTGCTCTGATGAATATACCTATGTCCTGAGAGTGGTACGTCAGGAGTGAATCG 1304  
DB |||||||  
QY 599 CTTCTGCTCTGATGAATATACCTATGTCCTGAGAGTGGTACGTCAGGAGTGAATCG 658  
DB |||||||  
QY 1305 GGAGCTGGGGGCTGGCAGGACCAAGCCTTAGGGGACCGTACCCGCTCCCTTACACCGA 1364  
DB |||||||  
QY 659 GGAGCTGGGGGCTGGCAGGACCAAGCCTTAGGGGACCGTACCCGCTCCCTTACACCGA 718  
DB |||||||  
QY 1365 CGGGGTTCTGATGAGCGGAGCGGTGTGCGCTGGTGGCCCATGGGAATACCCCGCAC 1424  
DB |||||||  
QY 719 CGCGGTTCTGATGAGCGGAGCGGTGTGCGCTGGTGGCCCATGGGAATACCCCGCAC 778  
DB |||||||  
QY 1425 CTTCTGCGGACCAACCGCTTCCGAGGTACACCTGCCCCAGGCGACGGAGTCTT-CC 1483  
DB |||||||  
QY 779 CTTCTGCGGACCAACCGCTTCCGAGGTACACCTGCCCCAGGCGAGGGCTTCCC 838  
DB |||||||  
QY 1484 CCTCTCTGGCTCCATCTCTGATGAGCCCCAAACATCTTCAAGCACCCAGAGAGTTCAACC 1543  
DB |||||||

DB 839 CCTCTCTGGCTCCATCTCTGATGACCCCACTCTTCCAGCACCCAGAGTTCAACC 898  
QY 1544 CAGAC 1548  
DB 899 CAGAC 903  
RESULT 9  
BU552425 915 bp mRNA linear EST 16-SEP-2002  
LOCUS AGENCOURT 10333108 NIH\_MGC\_40 Homo sapiens cDNA clone IMAGE:6576204  
5', mRNA sequence.  
BU552425  
BU552425.1 GI:22902697  
EST.  
Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 915)  
NIH-MGC http://mgi.nci.nih.gov/.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgabs-r@mail.nih.gov  
Tissue Procurement: DCTD/DTP  
cDNA Library Preparation: Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLCM2774 row: f column: 12  
High quality sequence stop: 754.

FEATURES  
Location/Qualifiers  
1..915  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6576204"  
/tissue\_type="carcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_40"  
/note="Organ: prostate; Vector: pOTB7; Site 1: XhoI;  
Site 2: EcoRI; cDNA made by oligo-dT priming.  
Directionally cloned into EcoRI/XhoI sites using the  
following 5' adaptor: GGCACGAG(G). Library constructed by  
Ling Hong in the laboratory of Gerald M. Rubin (University  
of California, Berkeley) using ZAP-cDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies).  
Note: this is a NIH\_MGC Library."

ORIGIN  
Query Match 28.3%; Score 833; DB 13; Length 915;  
Best Local Similarity 99.0%; Pred. No. 1.2e-95;  
Matches 858; Conservative 0; Mismatches 7; Indels 2; Gaps 2;

QY 1323 GGCACCAAGCCTAGGGGACCGTACCCGCTCCCTTACACGACGCGGTTCTGATGAGGC 1382  
DB 1 GGCACCAAGCCTAGGGGACCGTACCCGCTCCCTTACACGACGCGGTTCTGATGAGGC 60  
QY 1383 GCAGCGCTGTGGCGCTGGTGCCTATGGGAATACCCCGACCCCTCATCGGACCAACCG 1442  
DB 61 GCAGCGCTGTGGCGCTGGTGCCTATGGGAATACCCCGACCCCTCATCGGACCAACCG 120  
QY 1443 CTTTCGAGGGTACACCTGCGCCAGGGACGAGGTCTTCCCTCTCTGGTCCATCCT 1502  
DB 121 CTTTCGAGGGTACACCTGCGCCAGGGACGAGGTCTTCCCTCTCTGGTCCATCCT 180  
QY 1503 GCATGACCCCAACATCTTCAAGCACCCAGAGAGTTCAACCCAGACCGTTTCTGATGC 1562  
DB 181 GCATGACCCCAACATCTTCAAGCACCCAGAGAGTTCAACCCAGACCGTTTCTGATGC 240

FEATURES		Location/Qualifiers	
source		1. 877	
		/organism="Homo sapiens"	
		/mol_type="mRNA"	
		/db_xref="taxon:9606"	
		/clone="IMAGE:5446955"	
		/tissue_type="carcinoma, cell line"	
		/lab_host="DH10B (phage-resistant)"	
		/clone_lib="NIH_MGC_40"	
		/note="Organ: prostate; Vector: pOTB7; Site: 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."	
ORIGIN		Query Match 27.5%; Score 811; DB 12; Length 877;	
		Best Local Similarity 98.9%; Pred. No. 7.4e-93;	
		Matches 869; Conservative 0; Mismatches 5; Indels 5; Gaps 5;	
QY	1319	GCACGGCACCAAGCCTAGGGACCCCTACCGGCTCCCTTACACCGACGCGGTTCTGCATG	1378
DB	1	GCACGGCACCAAGCCTAGGGACCCCTACCGGCTCCCTTACACCGACGCGGTTCTGCATG	60
QY	1379	AGGCGACGGCTGCTGGCGTGGTCCCATGGGAATACCCGCGACCCCTCATCGGACCA	1438
DB	61	AGGCGACGGCTGCTGGCGTGGTCCCATGGGAATACCCGCGACCCCTCATCGGACCA	120
QY	1439	CCCGCTTCGAGGGTACACCCCTGCCCGGACGAGAGTCTTCCCTCTCTTGGCTCCA	1498
DB	121	CCCGCTTCGAGGGTACACCCCTGCCCGGACGAGAGTCTTCCCTCTCTTGGCTCCA	180
QY	1499	TCTCGATGACCCCAACATCTTCAAGCACCCAGAGAGTTCAACCCAGACCGTTTCTCTGG	1558
DB	181	TCTCGATGACCCCAACATCTTCAAGCACCCAGAGAGTTCAACCCAGACCGTTTCTCTGG	240
QY	1559	ATGCAATGAGCGGTTTCAAGAGCATGAGGCTTCTCTGCTCTCTTCTTAGGAGCGTG	1618
DB	241	ATGCAATGAGCGGTTTCAAGAGCATGAGGCTTCTCTGCTCTCTTCTTAGGAGCGTG	300
QY	1619	TCTGCTTGGAGAGGCGCTTGGCAAAAGGGAGCTCTTCTCTTCTTCAACCATCTTAC	1678
DB	301	TCTGCTTGGAGAGGCGCTTGGCAAAAGGGAGCTCTTCTCTTCTTCAACCATCTTAC	360
QY	1679	AAGCCTTCTCTGAGAGCCGCTGCGCGCGGACACCTGAGCCTCAAGCCACCGTCA	1738
DB	361	AAGCCTTCTCTGAGAGCCGCTGCGCGCGGACACCTGAGCCTCAAGCCACCGTCA	420
QY	1739	GTGGGCTTTTCAACATTTCCCGAGCTTCCAGCTGCAAGTGGTGGTGGTGGTGGTGGTGG	1798
DB	421	GTGGGCTTTTCAACATTTCCCGAGCTTCCAGCTGCAAGTGGTGGTGGTGGTGGTGGTGG	480
QY	1799	CCACCAACGACAGTGAAGAGGCAACTTGAAGTGGTGGTGGTGGTGGTGGTGGTGGTGG	1858
DB	481	CCACCAACGACAGTGAAGAGGCAACTTGAAGTGGTGGTGGTGGTGGTGGTGGTGGTGG	540
QY	1859	CTCCAGCCTCAACAGTGGGATGAGAGGTTAATGTCTCCAGAGTGTACACTGACGAGCA	1918
DB	541	CTCCAGCCTCAACAGTGGGATGAGAGGTTAATGTCTCCAGAGTGTACACTGACGAGCA	600
QY	1919	GCACATTTTACACGCTGAGTTGTTTCCGAGTGTCTGCCACGGCCACACGCTTCACTC	1977
DB	601	GCACATTTTACACGCTGAGTTGTTTCCGAGTGTCTGCCACGGCCACACGCTTCACTC	660
QY	1978	TTGACTCATGCTGCTAAGATGACAAACCGCACACCCATACACAACTTACAGGCGCACAA	2037
DB	661	TTGACTCATGCTGCTAAGATGACAAACCGCACACCCATACACAACTTACAGGCGCACAA	719
QY	2038	GCAACTGCTGGGTAGCTTTCCACAGACATAATAGTCCATCTCTCAATC-ACAAAGCAC	2096

```

Db      720  GCAACTGCTGGTAAAGTTTCCACAGACTTAAATATAGTCCATCTGCAATCAACAAGCAC 779
Qy      2097  ATAGCCAGGTAACCAACCAACTCCCTGATCTGCAGCCCACTGGAGTCTGGCTGT 2156
Db      780  ATAGCCAGGTAACCAACCAATCCCTGATCTGCAG--CCACAGCTGGAGTCTGGCTGT 838
Qy      2157  CACTTCAAGCCACAG-AAAACGGCCACACATGTTCCAC 2194
Db      839  CACCTTCACAAGCCACAGAAACGGGCACACATGTTCCC 877

RESULT 11
BQ958112 939 bp mRNA linear EST 21-AUG-2002
DEFINITION AGNCOURT_10034258 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6481143
5', mRNA sequence.
ACCESSION BQ958112
VERSION BQ958112.1 GI:22373590
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 939)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTP
cDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2662 row: a column: 16
High quality sequence stop: 636.
FEATURES
Location/Qualifiers
1..939
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6481143"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_40"
/note="Organ: prostate; Vector: pOTB7; Site 1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

ORIGIN
Query Match 27.4%; Score 807.8; DB 13; Length 939;
Best Local Similarity 96.8%; Pred. No. 1.8e-92;
Matches 877; Conservative 0; Mismatches 22; Indels 9; Gaps 5;

Qy      645  TGGGGTTTCTTCTTCCACGGGGAGCGGTGGAGCGAGCTGATCCAGCGGAGGCCCG 704
Db      1  TGGGGTTTCTTCTTCC-ACGGGAGCGGTGGAGCGAGCTGAGGAAGTTTACCATGCTTC 59

Qy      705  TCTGGGGACCTGGGCATGGGGAAGCGAAGCGAGGAGCTGATCCAGCGGAGGCCCG 764
Db      60  TCTGGGGACCTGGGCATGGGGAAGCGAAGCGAGGAGCTGATCCAGCGGAGGCCCG 119

Qy      765  GTGCTGTGGTGGACATTCAGGGGGACAGAAGACGCCCATTCGATCCCTCGCTGTGT 824
Db      120  GTGCTGTGGTGGACATTCAGGGGGACAGAAGACGCCCATTCGATCCCTCGCTGTGT 179

```

```

Qy      825  GGCCAGGCCACCTCCAAAGTAGTCTGCTCCTCTCTTTTGGCCTCCGCTTCTCCTATGA 884
Db      180  GGCCAGGCCACCTCCAAAGTAGTCTGCTCCTCTCTTTTGGCCTCCGCTTCTCCTATGA 239
Qy      885  GGATAAGAGATTCCAGGCCGTGGTCCGGGAGCTGGTGTACCTCTGCTGGAGTCAAGTTC 944
Db      240  GGATAAGAGATTCCAGGCCGTGGTCCGGGAGCTGGTGTACCTCTGCTGGAGTCAAGTTC 299
Qy      945  CCAGGGGGTCCAGACCTACGAGATGTTCTCTCTGGTTCCTGGGGCCCTTCCAGGCCCCA 1004
Db      300  CCAGGGGGTCCAGACCTACGAGATGTTCTCTCTGGTTCCTGGGGCCCTTCCAGGCCCCA 359
Qy      1005  CAAGCAGCTCTCCACACAGCTCAGCAGCTTGGTGTCTTCAAGTCCGGCAGGTGACAGA 1064
Db      360  CAAGCAGCTCTCTCCACACAGCTCAGCAGCTTGGTGTCTTCAAGTCCGGCAGGTGACAGA 419
Qy      1065  GCACAGGGGAACCTGGATGCTTCGGGCCCCGACGTCACCTTCGATGCTTCCTGCT 1124
Db      420  GCACAGGGGAACCTGGATGCTTCGGGCCCCGACGTCACCTTCGATGCTTCCTGCT 479
Qy      1125  GAAGATGGCACAGGAGGAACAAACCCAGGCACAGAAATTCACCAACAAGAACATGCTGAT 1184
Db      480  GAAGATGGCACAGGAGGAACAAACCCAGGCACAGAAATTCACCAACAAGAACATGCTGAT 539
Qy      1185  GACAGTCATTTATTTGCTGTTTCTGGGACATGACCGGTGAGCACCACGCTCGGTATAC 1244
Db      540  GACAGTCATTTATTTGCTGTTTCTGGGACATGACCGGTGAGCACCACGCTCGGTATAC 599
Qy      1245  CCTCTCTCTCTGATGAATACCTCATGTCCTCAAAAGTGGTACGTGAGGAGCTGAATCG 1304
Db      600  CCTCTCTCTCTGATGAATACCTCATGTCCTCAAAAGTGGTACGTGAGGAGCTGAATCG 659
Qy      1305  GGAGCTGGGGGTGGCCAGGCACCAAGCCCTAGGGGACCGTACCCGCCCTCCCTTACACCGA 1364
Db      660  GGAGCTGGGGGTGGCCAGGCACCAAGCCCTAGGGGACCGTACCCGCCCTCCCTTACACCGA 719
Qy      1365  CGCGGTCTGTCATGAGCGCAGCGGTGCTGGCGCTGGTGGCCATGGGAATACCCCGCAC 1424
Db      720  CGCGGNTCTGTCATGAGCGCAGCGGTGCTGGCGCTGGTGGCCATGGGAATACCCCGCAC 779
Qy      1425  CCTCATGGGACACCCCGC--TTCGAGGGTACACCTGCCCGCAGGC---ACGAGGTC 1479
Db      780  CCTCATGGGACACCCCGCTTTCCGAGGGTACACCTGCCCGCAGGCAGGAGTCTT 839
Qy      1480  TTCCCTCTCTTGGCTCCA--TCCTGCATG--ACCCCAACATCTTCAAGCACCCAGAGAG 1536
Db      840  TCCCCCTCTTGGCTCCA--TCCTGCATG--ACCCCAACATCTTCAAGCACCCAGAGAG 899
Qy      1537  TTCAACCC 1544
Db      900  TTTCAACC 907

RESULT 12
BQ047426
LOCUS BQ047426
DEFINITION 603628712F1 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:5457075 5',
mRNA sequence.
ACCESSION BQ047426
VERSION BQ047426.1 GI:16776693
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 857)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTP
cDNA Library Preparation: Ling Hong/Rubin Laboratory

```

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: LLCM1956 row: h column: 04  
 High quality sequence stop: 817.  
 Location/Qualifiers  
 1. 857  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5457075"  
 /tissue\_type="carcinoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH\_MGC\_40"  
 /note="Organ: prostate; Vector: pOTB7; Site: 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dr priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH\_MGC Library."

FEATURES  
 source  
 1. 857  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5457075"  
 /tissue\_type="carcinoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH\_MGC\_40"  
 /note="Organ: prostate; Vector: pOTB7; Site: 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dr priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH\_MGC Library."

ORIGIN  
 Query Match 27.2%; Score 801; DB 12; Length 857;  
 Best Local Similarity 98.5%; Pred. No. 1.4e-91;  
 Matches 840; Conservative 0; Mismatches 10; Indels 3; Gaps 3;  
 QY 641 GCCATGGGGTTTCTTCTCCACGGGGAGCGGTGGAGGCGCTGAGGAAGTTTACCATGC 700  
 Db 1 GCCATGGGGTTTCTTCTCCACGGGGAGCGGTGGAGGCGCTGAGGAAGTTTACCATGC 60  
 QY 701 TTGCTCTGCGGACCTGGCGATGGGAGCGAGAGCGGAGAGCTGATCCAGGGCGAGG 760  
 Db 61 TTGCTCTGCGGACCTGGCGATGGGAGCGAGAGCGGAGAGCTGATCCAGGGCGAGG 120  
 QY 761 CCGGTGTCTGTGGAGACATTCCAGGGGACAGAGGAGCGCCATTCGATCCCTCCCTGC 820  
 Db 121 CCGGTGTCTGTGGAGACATTCCA-GGACAGAGAGGAGCGCCATTCGATCCCTCCCTGC 179  
 QY 821 TGCTGCCAGCGACCTCCAGGTAGTGTCTCCCTCTCTTTGGCTCCGCTTCTCTCT 880  
 Db 180 TGCTGCCAGCGACCTCCAGGTAGTGTCTCCCTCTCTTTGGCTCCGCTTCTCTCT 239  
 QY 881 ATGAGGATAGGAGTTCCAGGCGGTGGTCCGGGACAGTGGTGGTACCTGTGGAGTCA 940  
 Db 240 ATGAGGATAGGAGTTCCAGGCGGTGGTCCGGGACAGTGGTGGTACCTGTGGAGTCA 299  
 QY 941 GCTCCAGGGGGTACAGCTACGAGATGTTCTCTGGTTCCTGGGCGCCCTGCCAGGCC 1000  
 Db 300 GCTCCAGGGGGTACAGCTACGAGATGTTCTCTGGTTCCTGGGCGCCCTGCCAGGCC 359  
 QY 1001 CCCACAAGCAGCTCTCCACCGTACAGACCTTGCTGCTTCCAGTCCGCGAGGTGC 1060  
 Db 360 CCCACAAGCAGCTCTCCACCGTACAGACCTTGCTGCTTCCAGTCCGCGAGGTGC 419  
 QY 1061 AGCAGCACCAGGGAAACCTGGATGCTTCGGGGCCCCGACGTCGTCGATGTCCTTCC 1120  
 Db 420 AGCAGCACCAGGGAAACCTGGATGCTTCGGGGCCCCGACGTCGTCGATGTCCTTCC 479  
 QY 1121 TGCTGAAGTGGCAGAGGAGGACAAACCCAGGACAGATTCACCAACAGAAATGC 1180  
 Db 480 TGCTGAAGTGGCAGAGGAGGACAAACCCAGGACAGATTCACCAACAGAAATGC 539  
 QY 1181 TGATGACAGTCATTATTGCTGTTGCTGGGACGATGAGGTCAGCACCACCGTCGGCT 1240  
 Db 540 TGATGACAGTCATTATTGCTGTTGCTGGGACGATGAGGTCAGCACCACCGTCGGCT 599  
 QY 1241 ATACCTCTCTGCTCCTGATGAATACCTCATGTCCAAAGTGGGTACGTGAGGAGCTGA 1300  
 Db 600 ATACCTCTCTGCTCCTGATGAATACCTCATGTCCAAAGTGGGTACGTGAGGAGCTGA 659

QY 1301 ATCGGGAGCTGGGGCTGGCCAGGCACCAAGCTTAGGGACCGTACCGCCTCCCTTACA 1360  
 Db 660 ATCGGGAGCTGGGGCTGGCCAGGCACCAAGCTTAGGGACCGTACCGCCTCCCTTACA 719  
 QY 1361 CCGACGGCGTTCTGCATGAGCGGCGGCTGTGGCGCTGTGGCCCATGGGAATACCC 1420  
 Db 720 CCGACGGCG-ICTGCATGAGCGGCGGTTGTGGCGCTGTGGCCCATGGGAATACCC 778  
 QY 1421 GCACCTCATGGGACCAACCGCTTCCGAGGGTACACCTGCCCGAGGCGACGGAGTCT 1480  
 Db 779 GGAACCTCATGGGACCAACCGCTTCCGAGGGTACA-CCTGCCCGAGGCGACGGAGTCT 837  
 QY 1481 TCCCTCTCTTGG 1493  
 Db 838 TCCCTCTCTGG 850

RESULT 13  
 BM044172  
 LOCUS 6036221408F1 NIH\_MGC\_40 Homo sapiens cDNA clone IMAGE:5447257 5',  
 DEFINITION mRNA sequence.  
 ACCESSION BM044172  
 VERSION BM044172.1 GI:16773439  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 804)  
 AUTHORS NIH-MGC <http://mgc.mci.nih.gov/>.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgabbs@mail.nih.gov](mailto:cgabbs@mail.nih.gov)  
 Tissue Procurement: DCTD/DTF  
 cDNA Library Preparation: Ling Hong/Rubin Laboratory  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: LLCM1930 row: o column: 02  
 High quality sequence stop: 801.  
 Location/Qualifiers  
 1. 804  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5447257"  
 /tissue\_type="carcinoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH\_MGC\_40"  
 /note="Organ: prostate; Vector: pOTB7; Site: 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dr priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH\_MGC Library."

FEATURES  
 source  
 1. 804  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5447257"  
 /tissue\_type="carcinoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH\_MGC\_40"  
 /note="Organ: prostate; Vector: pOTB7; Site: 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dr priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH\_MGC Library."

ORIGIN  
 Query Match 26.6%; Score 783.6; DB 12; Length 804;  
 Best Local Similarity 99.4%; Pred. No. 2.1e-89;  
 Matches 797; Conservative 0; Mismatches 4; Indels 1; Gaps 1;  
 QY 1319 GCCAGGACCAAGCTTAGGGACCGTACCGCCTCCCTTACACCGCGGTCTGCATG 1378  
 Db 1 GCCAGGACCAAGCTTAGGGACCGTACCGCCTCCCTTACACCGCGGTCTGCATG 60  
 QY 1379 AGCGGACGCGGTGTGGCGCTGGTCCCATGGGAATACCCCGCACCTCATGGGACCA 1438







GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 17, 2004, 19:05:01 ; Search time 5865 Seconds  
(without alignments)  
3724.619 Million cell updates/sec

Title: US-10-669-693-2  
Perfect score: 2615  
Sequence: 1 MEATGTWALLALALLLLT.....PPAFQLQVRPTDLHSTTQTR 504

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2\_l/USPTO.spool/US10669693/runat\_15092004\_101310\_5847/app\_query.fasta\_1.647  
-DB=GenEmbl -OPMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10669693 @CGN 1 1 3731 @runat\_15092004\_101310\_5847 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSFBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DLEXT=7

Database :

GenEmbl:\*  
1: gb\_ba:\*  
2: gb\_htg:\*  
3: gb\_in:\*  
4: gb\_cm:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vi:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_on:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sts:\*  
28: em\_un:\*

29: em\_vi:\*  
30: em\_htg\_hum:\*  
31: em\_htg\_inv:\*  
32: em\_htg\_other:\*  
33: em\_htg\_mus:\*  
34: em\_htg\_pln:\*  
35: em\_htg\_rtd:\*  
36: em\_htg\_mam:\*  
37: em\_htg\_vrt:\*  
38: em\_sy:\*  
39: em\_htg\_hum:\*  
40: em\_htg\_mus:\*  
41: em\_htg\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match %	Length	DB ID	Description
1	2615	100.0	1515	6	AX591678 Sequence
2	2615	100.0	2261	6	AX780433 Sequence
3	2615	100.0	2261	9	AF335278 Homo sapi
4	2615	100.0	2604	6	AX430320 Sequence
5	2615	100.0	2619	9	BC033691 Homo sapi
6	2615	100.0	2620	6	AX327331 Sequence
7	2615	100.0	2623	6	AX463930 Sequence
8	2615	100.0	2626	9	AY386603 Homo sapi
9	2610	99.8	2604	6	AX552230 Sequence
10	2527	96.6	2513	6	AX876954 Sequence
11	2527	96.6	2513	6	BD156406 Sequence
12	2527	96.6	2513	9	AK027605 Homo sapi
13	1628.5	62.3	2114	6	AX552231 Sequence
14	1313	50.2	2026	4	RABF450BX
15	1313	50.2	2028	4	S64259
16	1301	49.8	2045	4	RABCYCP2B
17	1301	49.8	2743	4	AB052256 Sus scrofa
18	1300	49.7	2065	4	RABCYCP2C
19	1300	49.7	2081	10	BC034202 Mus muscu
20	1299.5	49.7	1795	10	AF128849 Mus muscu
21	1299.5	49.7	1901	10	BC060973 Mus muscu
22	1299.5	49.7	2079	4	RABCYCP2A
23	1299	49.7	1802	4	RABP450CA
24	1289	49.3	1567	10	RATCYP450
25	1289	49.3	1567	10	RATCYP45Z
26	1286.5	49.2	2004	6	AX477702 Sequence
27	1280.5	49.0	1822	10	MUSTH16AB
28	1262	48.3	2642	4	DOGPA591IB
29	1261.5	48.2	1717	10	D86952
30	1257	48.1	1480	6	E63778
31	1255	48.0	1988	10	AY250712 Mesocricetu
32	1245.5	47.6	1853	10	L81171 Mus musculu
33	1245.5	47.6	1886	10	BC058222 Mus muscu
34	1244	47.6	1853	10	MUSTH16A
35	1243.5	47.6	1857	10	RATCYT45XX
36	1242.5	47.5	1781	4	AB052255 Sus scrofa
37	1239	47.4	3045	6	AX018551 Sequence
38	1239	47.4	3045	6	AX018627 Sequence
39	1239	47.4	3045	6	BD136285 Enhanceme
40	1238	47.3	1476	6	E10616 Human CDNA
41	1238	47.3	1476	6	E10861 CDNA encodi
42	1238	47.3	1476	6	AR399329 Sequence
43	1238	47.3	3045	6	AX332736 Sequence
44	1238	47.3	3045	6	AX409672 Sequence
45	1238	47.3	3045	9	M29874 Human cytoc

ALIGNMENTS

RESULT 1

AX591678  
LOCUS AX591678 1515 bp DNA linear PAT 27-JAN-2003  
DEFINITION Sequence 39 from Patent WO246409.  
ACCESSION AX591678  
VERSION AX591678.1 GI:27950054  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1  
Guo,X., Li,L., Patturajan,M., Shimkets,R.A., Casman,S.J.,  
Malyanar,U.M., Tchernev,V.T., Vernet,C.A., Spytek,K.A.,  
Shenoy,S.G., Alsobrook,J.P., Edinger,S., Peyman,J.A., Stone,D.J.,  
Ellerman,K., Gangolli,E.A., Boldog,F.L., Colman,S.D., Eisen,A.J.,  
Liu,X., Padigaru,M., Spaderna,S.K. and Zerhusen,B.D.  
Proteins and nucleic acids encoding same  
Patent: WO 0246409-A 39 13-JUN-2002;  
Curagen Corporation (US)  
FEATURES  
Location/Qualifiers  
1..1515  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Alignment Scores:  
Pred. No.: 2,69e-201 Length: 1515  
Score: 2615.00 Matches: 504  
Percent Similarity: 100.00% Conservatives: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0  
US-10-669-693-2 (1-504) x AX591678 (1-1515)  
QY 1 MetGluAlaThrGlyThrTrpAlaLeuLeuLeuAlaLeuLeuLeuLeuLeuThr 20  
Db 1 ATGGAGCGACCGGACCTGGGGCTGCTGGCGCTGGCGCTGGCTGGCTGGCG 60  
QY 21 LeuAlaLeuSerGlyThrArgAlaArgGlyHisLeuProGlyProThrProLeuPro 40  
Db 61 CTGGCGCTGTCGGGACCGAGGCGCCGAGGCGACCTGCCCGCGGCGCCGCGCTACCA 120  
QY 41 LeuLeuGlyAsnLeuLeuGlnLeuArgProGlyValaLeuLeuTyrSerGlyLeuMetArgLeu 60  
Db 121 CTGCTGGGAAACCTCTGAGCTACGGCCGGGCGCTGTATTTCAGGGCTCATGGCGCTG 180  
QY 61 SerLysLysTyrGlyProValPheThrIleTyrLeuGlyProThrArgProValVal 80  
Db 181 AGTAAGAAGTACGACCGGTGTTACCATCTACCTGGGACCGTGGCGGCTGGTGGTGC 240  
QY 81 LeuValGlyGlnGlnAlaValArgGluAlaLeuLeuGlyGlyGlnAlaGluGluPheSerGly 100  
Db 241 CTGGTGGGCGAGAGGCTGTGGGAGGCGCCCTGGAGGTACGAGTGGAGGATTCAGCGGC 300  
QY 101 ArgGlyThrValAlaMetLeuGluGlyThrPheAspGlyHisGlyValPhePheSerAsn 120  
Db 301 CGGGGAACCTAGCATGCTGGAAGGACATTTGATGGCCATGGGGTTTCTTCTTCCAAC 360  
QY 121 GlyGluArgTrpArgGlnLeuArgLysPheThrMetLeuAlaLeuArgAspLeuGlyMet 140  
Db 361 GGGGAGCGGTGGGAGGAGCTGAGGAAGTTTACCATGCTTGCCTGCGGGACCTGGGCATG 420  
QY 141 GlyLysArgGluGlyGluGluLeuIleGlnAlaGluAlaArgCysLeuValGluThrPhe 160  
Db 421 GGGAGCGAGAGGCGAGGAGCTGATCCAGCGGAGGCGCGCGGTGCTGTGGTGGACATTC 480  
QY 161 GlnGlyThrGluGlyArgProPheAspProSerLeuLeuLeuAlaGlnAlaThrSerAsn 180  
Db 481 CAGGGGACAGAGGAGCGCCATTTCGATCCCTCCCTGCTGTGGCCAGGCCACCTCAAC 540  
QY 181 ValValCysSerLeuLeuPhePheGlyLeuArgPheSerTyrGluAspLysGluPheGlnAla 200

Db 541 GTAGTCTGCTCCCTCTCTTTGGCCTCGCTTCTCTATGAGGATAAGGAGTTCAGGCC 600  
QY 201 ValValArgAlaAlaGlyGlyThrLeuLeuGlyValSerSerGlnGlyGlyGlnThrTyr 220  
Db 601 GTGGTCCGGGCGAGCTGGTGTACCTGCTGGAGTTCAGCTCCCGGGGGGTGACAGCTTAC 660  
QY 221 GluMetPheSerTrpPheLeuArgProLeuProGlyProHisLysGlnLeuLeuHisHis 240  
Db 661 GAGATGTTCTCTGGTTCCTGGGCCCTCTGCAGGCCCCCAAGCAGCAGCTCTCTCCACCAC 720  
QY 241 ValSerThrLeuAlaAlaPheThrValArgGlnValGlnGlnHisGlnGlyAsnLeuAsp 260  
Db 721 GTGAGCAGCTTGGTGGCTTCCAGATCGCGCAGGTGCAGCAGCACACGGGAAACCTGGAT 780  
QY 261 AlaSerGlyProAlaArgAspLeuValAspAlaPheLeuLeuLysMetAlaGlnGluGlu 280  
Db 781 GCTTCGGGCCCGCACGCTGACCTTGTGATGCTTCTTCGTGAGATGTCACAGGAGNA 840  
QY 281 GlnAsnProGlyThrGluPheThrAsnLysAsnMetLeuMetThrValIleTyrLeuLeu 300  
Db 841 CAAACCCAGGACAGAAATTCACCAACAAGACATGCTGATGACAGTCATTTATTTCGTG 900  
QY 301 PheAlaGlyThrMetThrValSerThrThrValGlyTyrThrLeuLeuLeuMetLys 320  
Db 901 TTTGCTGGGAGCATGACGCTCAGCACCGCTGGCTATACCTCTCTCTCTGATGAAA 960  
QY 321 TyrProHisValGlnLysTrpValArgGluGluLeuAsnArgGluLeuGlyAlaGlyGln 340  
Db 961 TACCCTCATGTCCAAAGTGGGTACGTGAGGAGCTGAATCGGAGCTGGGGCTGGCCAG 1020  
QY 341 AlaProSerLeuGlyAspArgThrArgLeuProTyrThrAspAlaValLeuHisGluAla 360  
Db 1021 GCCCAAGCCTAGGGACCGTACCCGCTCCCTTACACCGACGCGCTTCTGTCATGAGCG 1080  
QY 361 GlnArgLeuLeuAlaLeuValProMetGlyIleProArgThrLeuMetArgThrThrArg 380  
Db 1081 CAGCGCTGCTGGCGTGGTGGCCATGGGAATACCCCGCACCTCATGCGGACCCCGC 1140  
QY 381 PheArgGlyTyrThrLeuProGlnGlyThrGluValPheProLeuLeuGlySerIleLeu 400  
Db 1141 TTTCCGAGGTATACACCTGCCCGAGGCGACGGAGGTCTTCCCTCTCTTGGCTCCATCT 1200  
QY 401 HisAspProAsnIlePheLysHisProGluGluPheAsnProAspArgPheLeuAspAla 420  
Db 1201 CATGACCCCAACATCTTCAGCACCCAGAGAGTTCAACCCAGACCGTTTCTCTGATGCA 1260  
QY 421 AspGlyArgPheArgLysHisGluAlaPheLeuProPheSerLeuGlyLysArgValCys 440  
Db 1261 GATGACGCTTCAGGAAGCATGAGGCGTTCCTGCCCTTCTCTTAGGAGACGCTGTCTG 1320  
QY 441 LeuGlyGluGlyLeuAlaLysAlaGluLeuPhePhePheThrThrIleLeuGlnAla 460  
Db 1321 CTTTGGAGAGGCTGGCAAAAGCGAGCTCTTCTCTTCTTACCCACCATCTTACAGCC 1380  
QY 461 PheSerLeuGluSerProCysProProAspThrLeuSerLeuLysProThrValSerGly 480  
Db 1381 TTTCTCTCGAGAGCCCGTGGCCCGCGGACACCTCAGCCTCAAGCCACCGTCAGTGGC 1440  
QY 481 LeuPheAsnIleProProAlaPheGlnLeuGlnValArgProThrAspLeuHisSerThr 500  
Db 1441 CTTTTCAACATTCGCCCGAGCTTCCAGCTGCAAGTCGCTCCCTGACCTTCTACTCCACC 1500  
QY 501 ThrGlnThrArg 504  
Db 1501 ACGAGACCCAGA 1512  
RESULT 2  
AX780433 2261 bp DNA linear PAT 14-JUL-2003  
LOCUS AX780433  
DEFINITION Sequence 2590 from Patent WO03039443.  
ACCESSION AX780433  
VERSION AX780433.1 GI:32697427



AUTHORS	Rylander,T., Neve,E.P., Ingelman-Sundberg,M. and Oscarson,M.
TITLE	Identification and tissue distribution of the novel human cytochrome P450 2S1 (CYP2S1)
JOURNAL	Biochem. Biophys. Res. Commun. 281 (2), 529-535 (2001)
MEDLINE	21092856
PUBMED	11181079
REFERENCE	2 (bases 1 to 2261)
AUTHORS	Rylander,T., Neve,E., Ingelman-Sundberg,M. and Oscarson,M.
TITLE	Direct Submission
JOURNAL	Submitted (10-JAN-2001) Institute of Environmental Medicine, Molecular Toxicology, Karolinska Institute, Nobelsv. 13, Stockholm 17177, Sweden
FEATURES	Location/Qualifiers
source	1..2261
	/organism="Homo sapiens"
	/mol_type="mRNA"
	/db_xref="taxon:9606"
	/chromosome="19"
gene	1..2261
	/gene="CYP2S1"
CDS	1..1515
	/gene="CYP2S1"
	/codon_start=1
	/product="cytochrome P450 2S1"
	/protein_id="AAK13498.1"
	/db_xref="GI:13161184"
	/translation="MEATGTWALLALALALLLALSGTRARGHLPPGPTPLPLLGN LLOLRPALYSGLMLSKYGFVFTIYLGPMRPVVVLVGOEAVREALQOAEBSFSGRG TVMLEGTFDGHGVFPSNGERWOLKRTMLALRDLGMKREGEELIQAEARCLIVET QTEGRFPDPSLLLAQATSNVVCSSLFGLRFSYEDKEFOAVRAAGGTLIGVSSOGGQ TYEMSFLEPFLPGQQLLHHVSTLAAPTQVQOQHOGNLDASGPARDLVADFLMKM AQEQNPQTEFTKMLMTVIYLLFAGTMTVSTVGYITLLLMKTPHVKQVRELNRL ELGAGQAPSLGDRTPDYTDVILHEAQRLALVPMGIPRTLLMRTFRFGYITLPQSTEV PFLIGLILHDPNIFKHPBEFNDPDLADGRFRKHEAFLPFLSLGRKRVCLGELAKAEI PLFFTIIQAFLSESPCPDITLSLQPTVSGLFNIPPAFQLOVRPTDLHSTTQTR"
ORIGIN	
Alignment Scores:	
Pred. No.:	4, 426-201
Score:	2615.00
Length:	2261
Matches:	504
Percent Similarity:	100.00%
Conservative:	0
Best local Similarity:	100.00%
Mismatches:	0
Query Match:	100.00%
Indels:	0
Gaps:	9
DB:	
US-10-669-693-2 (1-504) x AF335278 (1-2261)	
Qy	1 MetGluAlaThrGlyThrTrpAlaLeuLeuAlaLeuAlaLeuLeuLeuLeuLeuThr 20
Db	1 ATGGAGCGCACCGCACCTGGCGCTGCTGGCGCTGGCGCTGCTCTCTGCTGCTGACG 60
Qy	21 LeuAlaLeuSerGlyThrArgAlaArgGlyHisLeuProProGlyProThrProLeuPro 40
Db	61 CTGGCGCTGCTCGGACCGAGGCGCGGAGCCACCTGCCCGCGGCGCCACGCGCTACCA 120
Qy	41 LeuLeuGlyAsnLeuLeuGlnLeuArgProGlyAlaLeuThrSerGlyLeuMetArgLeu 60
Db	121 CTGCTGGGAAACCTCTCGAGCTACGCGCGCGCGCGCTGTATTTCAGCGGCTCATGCGGCTG 180
Qy	61 SerLysIleThrGlyProValPheThrIleTyrLeuGlyProTrpArgProValValVal 80
Db	181 AGTAAGAAGTACCGACCGGGTTCACCACTACTGGGACCTGGCGGCTGTGGTGGTC 240
Qy	81 LeuValGlyGlnGluAlaValArgGluAlaLeuGlyGlyGlnAlaGluGluPheSerGly 100
Db	241 CTGGTTGGGCAGAGGCTGTGGGAGGCGCTGGAGGCTCAGGAGTTCAGGAGTTCAGCGGC 300
Qy	101 ArgGlyThrValAlaMetLeuGluGlyThrPheAspGlyHisGlyValPhePheSerAsn 120
Db	301 CGGGGAACCGTAGCGATGCTGGGAAGGACTTTTATGGCCCATGGGGTTTCTTCTCCAAC 360
Qy	121 GlyGluArgTrpArgGlnLeuArgLysPheThrMetLeuAlaLeuArgAspLeuGlyMet 140

Db	361 GGGGAGCGGTGGAGGAGCTGAGAGAGTTTACCATGCTGTGCTCGGGACCTGGGCGATG 420
Qy	141 GlyIysArgGluGlyGluGluLeuLeuLeuAlaGluAlaArgCysLeuValGluThrPhe 160
Db	421 GGGGAAGCAGAGCGGAGGAGCTGATCAGGCGGAGGCGCGGTGTCTGTGGACATTC 480
Qy	161 GlnGlyThrGluGlyArgProPheAspProSerLeuLeuLeuAlaGlnAlaThrSerAsn 180
Db	481 CAGGGGACAGAGACGCCCATTCGATCCCTCCCTGCTGTGGCCAGGCCACCTCCAAC 540
Qy	181 ValValCysSerLeuLeuPheGlyLeuArgPheSerTyrGluAspLysGluPheGlnAla 200
Db	541 GTAGTCTGCTCCCTCTCTTTGGCTCGCTCTCTATAGAGATAAGAGATTCAGGCC 600
Qy	201 ValValArgAlaAlaGlyThrLeuLeuGlyValSerSerGlnGlyGlyGlnThrTyr 220
Db	601 GTGGTCCGGGAGCTGGTGTTACCTCTGCTGGAGTACGATCCAGGGGGGTTCAGACCTAC 660
Qy	221 GluMetPheSerTrpPheLeuArgProLeuProGlyProHisLysGlnLeuLeuHisHis 240
Db	661 GAGATGTTCTCTGTTCTCTCGGCCCTCGCAGGCCCCCAAGACAGCTCTCTCCACCAC 720
Qy	241 ValSerThrLeuAlaAlaPheThrValArgGlnValGlnGlnHisGlnGlyAsnLeuAsp 260
Db	721 GTCAGACCTTTGGTGGCTTTCACAGTCCGGCAGGTGCAGCACACAGGGGAACCTGGAT 780
Qy	261 AlaSerGlyProAlaAlaAspLeuValAspAlaPheLeuLeuLysMetAlaGlnGlu 280
Db	781 GCTTCGGGCCCCGACAGTACCTTGTGATGCCTTCCTGCTGAAGATGGCAGGAGGAA 840
Qy	281 GlnAsnProGlyThrGluPheThrAsnLysAsnMetLeuMetThrValIleTyrLeuLeu 300
Db	841 CAAAAACCCAGGACAGATTCACCAACAAGAACATGCTGATGACAGTCATTTATTGTGTG 900
Qy	301 PheAlaGlyThrMetThrValSerThrThrValGlyTyrThrLeuLeuLeuMetLys 320
Db	901 TTTGCTGGGACGATGACGCTCAGCACCGGTGCGCTATACCTCTCTCTCTCTGATGAAA 960
Qy	321 TyrProHisValGlnLysTrpValArgGluGluLeuAsnArgGluLeuGlyAlaGlyGln 340
Db	961 TACCCTCATGTCCTCAAAAGTGGGTACGTGAGGAGCTGATCGGAGCTGGGGCTGGCCAG 1020
Qy	341 AlaProSerLeuGlyAspArgThrArgLeuProTyrThrAspAlaValLeuHisGluAla 360
Db	1021 GCACCAAGCTTAGGGGACCGTACCCTCCCTTACACGACGCGGTTCGTGATGAGGCG 1080
Qy	361 GlnArgLeuLeuAlaLeuValProMetGlyIleLeuProArgThrLeuMetArgThrArg 380
Db	1081 CAGCGCTGCTGGCGCTGGTGCCTATGGGAATACCCCGCACCTCATGGCGACACCCCGC 1140
Qy	381 PheArgGlyTyrThrLeuProGlnGlyThrGluValPheProLeuLeuGlySerIleLeu 400
Db	1141 TTCCGAGGGTACACCTGCTGCCCGGACGAGGAGTCTTCCCTCTCTTGGCTCCATCTG 1200
Qy	401 HisAspProAsnIlePheLysHisProGluGluPheAsnProAspArgPheLeuAspAla 420
Db	1201 CATGACCCCAACATCTTCAAGCACCCAGAAAGAGTTCAACCCAGACCGTTTCTTGGATGCA 1260
Qy	421 AspGlyArgPheArgLysHisGluAlaPheLeuProPheSerLeuGlyLysArgValCys 440
Db	1261 GATGGACGGTTCAGGAAGCATGAGGCGTTCCTGCCCTTCTCTTAGGGAAGCGTGTCTGC 1320
Qy	441 LeuGlyGluGluLeuAlaLysAlaGluLeuPhePheThrThrIleLeuGlnAla 460
Db	1321 CTTGGAGGGCTGGAAGAGGAGGAGCTTCTCTTCTTCCACCATCTTACAGGCC 1380
Qy	461 PheSerLeuGluSerProCysProAspThrLeuSerLeuLysProThrValSerGly 480
Db	1381 TTCTCTCTCGAGAGCGGTGCCCGCGGACACCTGAGCCTCAAGGCCACCGTCAGTGGC 1440
Qy	481 LeuPheAsnIleProProAlaPheGlnLeuGlnValArgProThrAspLeuHisSerThr 500
Db	1441 CTTTTCACATTCGCCAGCCTTCCAGCTGCAAGTCCGTCCTCCACTGACCTTCACTCCACC 1500

QY	501	ThrGlnThrArg	504	Db	520	CAGGGGACAGAAAGGACGCCCAATTCGATCCCTCCCTGCTGCTGGCCAGGCCACCTCCAAAC	579
Db	1501	ACGCAGACCAGA	1512	QY	181	ValValCysSerLeuLeuPheGlyLeuArgPheSerTyrGluAspLysGluPheGlnAla	200
RESULT 4				Db	580	GTAGTCTGCTCCTCTCTTTGGCCCTCCCTCTCTATGAGATGAAGAGATTTCCAGGCC	639
LOCUS	AX430320			QY	201	ValValArgAlaAlaGlyGlyThrLeuLeuGlyValSerSerGlnGlyGlyGlnThrTyr	220
DEFINITION	Sequence 102 from Patent WO0240715.			Db	640	GTGTCGCGGACAGCTGGTACCTGTCTGGAGTCAGCTCCAGGGGGGTTCAGACCTAC	699
ACCESSION	AX430320			QY	221	GluMetPheSerTyrPheLeuArgProLeuProGlyProHisGlnLeuLeuHisHis	240
VERSION	AX430320.1	GI:21655684		Db	700	GAGATGTTCTCTGCTGCTTCTGCGCCCTTCAGAGCCGCCCAAGCAGCTCTCCACCCAC	759
KEYWORDS				QY	241	ValSerThrLeuAlaAlaPheThrValArgGlnValGlnGlnHisGlnGlyAsnLeuAsp	260
SOURCE	Homo sapiens (human)			Db	760	GTACAGCACCTTGGCTGCTTTCACAGTCCGGCAGGTGACAGCACCAGGGAACTGGAT	819
ORGANISM	Homo sapiens			QY	261	AlaSerGlyProAlaArgAspLeuValAspAlaPheLeuLeuLysMetAlaGlnGlu	280
REFERENCE	1			Db	820	GCTTGGGCCCCGACGTCAGCTTGTGATGCTTCTCTGCTGAAGATGSCACAGAGGAA	879
AUTHORS	Chalup,M.S., Altus,C.M., Lincoln,S.E., Dufour,G.E. and Jackson,S.			QY	281	GlnAsnProGlyThrGluPheThrAsnLysAsnMetLeuMetThrValIleTyrLeuLeu	300
TITLE	Molecules for disease detection and treatment			Db	880	CAAAACCCAGGCACAGAAATTCACCAACAAGAACATGTGTGATGACAGTATTATTGCTG	939
JOURNAL	Patent: WO 0240715-A 102 23-MAY-2002;			QY	301	PheAlaGlyThrMetThrValSerThrThrValGlyTyrThrLeuLeuLeuLeuMetLys	320
INCYTE GENOMICS INC (US)				Db	940	TTTGTGGGACGATGACGGTCAGCCAGCACCGTTCGGCTATACCTCTCTCTGATGAAA	999
FEATURES	Location/Qualifiers			QY	321	TyrProHisValGlnLysTyrValArgGluGluLeuAsnArgGluLeuGlyAlaGlyGln	340
source	1..2604			Db	1000	TACCTTCATGTCCAAAAGTGGTACGTGAGGAGTGAATCGGAGCTGGGGCTGGCCAG	1059
	/organism="Homo sapiens"			QY	341	AlaProSerLeuGlyAspArgThrArgLeuProTyrThrAspAlaValLeuHisGluAla	360
	/mol_type="unassigned DNA"			Db	1060	GCACCAAGCTTAGGGGACCGTACCCGCTCTCTTACACCGACGCGGTCTTGCATGAGCG	1119
	/db_xref="taxon:9606"			QY	361	GlnArgLeuLeuAlaLeuValProMetGlyIleProArgThrLeuMetArgThrThrArg	380
	/note="Incyte ID No: LG:337818.2:2000SEP08"			Db	1120	CAGCGGCTGTGGCGCTGGTGCCCATGGGAATACCCCGCACCTTCATGGCGACCCCGC	1179
Alignment Scores:				QY	381	PheArgGlyTyrThrLeuProGlnGlyThrGluValPheProLeuLeuGlySerIleLeu	400
Pred. No.:	5,26e-201	Length:	2604	Db	1180	TTCCGAGGGTACACCTTCGCCCGCAGGGCACGAGGTCTTCCCTCTCTTGCCTCATCTG	1239
Score:	2615.00	Matches:	504	QY	401	HisAspProAsnIlePheLysHisProGluGluPheAsnProAspArgPheLeuAspAla	420
Percent Similarity:	100.00%	Conservative:	0	Db	1240	CATGACCCCAACATCTTCAAGCACCCAGAGAGTTCAACCCAGACCGCTTCTCTGGATGA	1299
Best Local Similarity:	100.00%	Mismatches:	0	QY	421	AspGlyArgPheArgLysHisGluAlaPheLeuProPheSerLeuGlyLysArgValCys	440
Query Match:	100.00%	Indels:	0	Db	1300	GATGGACGGTTTCAGGAAGCATGAGGGGCTTCTGCTCCCTCTCTTAGGAAGCGGTCTGC	1359
DB:	6	Gaps:	0	QY	441	LeuGlyGlyGlyLeuAlaLysAlaGluLeuPhePhePheThrIleLeuGlnAla	460
US-10-669-693-2 (1-504) x AX430320 (1-2604)				Db	1360	CTTGGAGAGGGCTGGCAAAAGCGAGCTTCTCTCTTTCACCAACCATCTCTACAGCC	1419
QY	1	MetGluAlaThrGlyThrTrpAlaLeuLeuLeuAlaLeuAlaLeuLeuLeuThr	20	QY	461	PheSerLeuGluSerProCysProProAspThrLeuSerLeuLysProThrValSerGly	480
Db	40	ATGAGGCGACCGGCACCTGGCGCTGCTGTGGCGTGGCGTGTCTCTGCTGACG	99	Db	1420	TTCTCTCTGGAGGCCCGTGGCCCGCGGACACCTTGAGCTTCAAGCCCAACCGTCAGTGC	1479
QY	21	LeuAlaLeuSerGlyThrArgAlaArgGlyHisLeuProGlyProThrProLeuPro	40	QY	481	LeuPheAsnIleProProAlaPheGlnLeuGlnValArgProThrAspLeuHisSerThr	500
Db	100	CTGGCGCTGTCCGGGACACGAGGCCGAGGCCACCTGTGCCCCCGGGCCACCGCGCTACCA	159	Db	1480	CTTTTCAACATTTCCCCAGCCTTCCAGCTGCAAGTCCGTCCTCCACTGACCTTCACTCCACC	1539
QY	41	LeuLeuGlyAsnLeuLeuGlnLeuArgProGlyAlaLeuTyrSerGlyLeuMetArgLeu	60	QY	501	ThrGlnThrArg	504
Db	160	CTGCTGGGAACCTCTCTGCAGCTACGCGCCGCGGCGCTGTATTTCAGGGCTCATGCGCTG	219	Db	1540	ACGCAGACCAGA	1551
QY	61	SerLysLysTyrGlyProValPheThrIleTyrLeuGlyProThrArgProValValVal	80	RESULT 5			
Db	220	AGTAAGAAGTACGACCGGTGTTCACCACTACCTGAGACCCCTGGCGGCGCTGGTGGTC	279	BC033691			
QY	81	LeuValGlyGlnGluAlaValArgGluAlaLeuGlyGlyGlnAlaGluGluPheSerGly	100	LOCUS			
Db	280	CTGGTTGGGACGAGGCTGTGCGGAGGCCCTCGGAGGTACGGCTGAGGAGTTTTCCTTCCAAC	339				
QY	101	ArgGlyThrValAlaMetLeuGluGlyThrPheAspGlyHisGlyValPhePheSerAsn	120				
Db	340	CGGGGAAACCGTAGCGAGTGTGGAAAGGACTTTTGTATGGCATGGGGTTCCTTCTCCAAC	399				
QY	121	GlyGluArgTrpArgGlnLeuArgLysPheThrMetLeuAlaLeuArgAspLeuGlyMet	140				
Db	400	GGGAGCGGTGGAGGACGCTGAGGAAGTTTACCATGCTTGTCTGCGGACCTGGGCATG	459				
QY	141	GlyLysArgGluGlyGluLeuIleGlnAlaGluAlaArgCysLeuValGluThrPhe	160				
Db	460	GGGAAGCGAAGAGCGGAGGAGCTGATCCAGCGGAGGCCCGGTCTGTGTGGAGACATTC	519				
QY	161	GlnGlyThrGluGlyArgProPheAspProSerLeuLeuLeuAlaGlnAlaThrSerAsn	180				

DEFINITION Homo sapiens cytochrome P450, family 2, subfamily S, polypeptide 1, mRNA (cdna clone MGC:44853 IMAGE:5212609), complete cds.

ACCESSION BC033691  
VERSION BC033691.1 GI:21707057  
KEYWORDS MGC.  
SOURCE Homo sapiens  
ORGANISM Homo sapiens (human)

REFERENCE 1 (bases 1 to 2619)  
AUTHORS Strausberg, R.D., Collins, F.S., Wagner, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L.H., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Datchenko, L., Soares, K.B., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, J.E., Brownstein, M.J., Uesdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Rulyk, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettner, M., Madan, A., Young, A.C., Rodrigues, S., Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickinson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalish, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.

TITLE Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

MEDLINE 22388257  
PUBMED 12477932

REFERENCE 2 (bases 1 to 2619)  
AUTHORS Strausberg, R.

TITLE Direct Submission  
JOURNAL Submitted (02-JUL-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
COMMENT Contact: [cgabbs@mail.nih.gov](mailto:cgabbs@mail.nih.gov)  
Tissue Procurement: Life Technologies, Inc.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland;  
Web site: <http://www.nisc.nih.gov/>  
Contact: [nisc.mgc@nih.gov](mailto:nisc.mgc@nih.gov)  
Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R., Maduro, Q.L., Masiello, C., Maki, B., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Stantirpop, S., Thomas, P.J., Touchman, J.W., Tsourgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.

FEATURES source  
1..2619  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="MGC:44853 IMAGE:5212609"  
/tissue type="Blood, adult leukocytes"  
/clone\_lib="NIH\_MGC\_118"

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAC Plate: 68 Row: 0 Column: 11  
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 20522238.

Location/Qualifiers  
1..2619

101 ArgGlyThrValAlaMetLeuGluGlyHisGlyValPhePheSerAsn 120  
318 CGGGAAACCGTAGCATGCTGGAAGGACTTTTGTATGCCATGGGGTTTCTTCTCCAC 377  
121 GlyClnuArgTrpArgGlnLeuArgGlyPheThrMetLeuAlaLeuArgAspLeuGlyMet 140  
378 GGGAGCGGTGGAGGACCTGAGAAAGTTTACCATGTTGCTCTCGCGGACCTGGGCATG 437  
141 GlyLysArgGluGlyGluGluLeuIleGlnAlaGluAlaArgCysLeuValGluThrPhe 160  
438 GGGAGCGAGAAAGCGAGGAGCTGATCCAGCGGAGGCCCGCTGTCTGTGGAGACATTC 497  
161 GlnGlyThrGluGlyArgPropPheAspProSerLeuLeuAlaGlnAlaThrSerAsn 180

gene

CDS

/lab\_host="DH10B"  
/note="vector: pCMV-SPORT6"  
1..2619  
/gene="CYP2S1"  
/db\_xref="LocusID:29785"  
18..1532  
/codon\_start=1  
/product="cytochrome P450, family 2, subfamily S, polypeptide 1"  
/protein\_id="AAH33691.1"  
/db\_xref="GI:21707058"  
/db\_xref="LocusID:29785"  
/translation="MEATGTWALLALALLLLLTALSGTRARGHLPGPTPLPGLN LLQRPALYSGLMRLSKYGVFTYILGPRVVLVVGQAVREALGGQAEFFGRG TVMLEGTFDGHVFFSNGRWRORLRFKFTMLALDRMGKREBELIQAEARCLVET QCTGRPDPSLLAQATSNVCSLLFLGRFSYEDKEFOAVRAAGGTLGLVSSQGG TYEWSFLRPLPGLHQHLLHVSTLANFTVQVQHOGLNDASGPARDLVDAFLKM AOEQNPGTEFTNKMMLTVLLFAGMTNRTVTVGYTLLLLMKYPHVQKAVREINR ELGAGQPSLGDRTLPYDAVLEAQLALVPMGIPRTLMRTTRFRGYTLPGQTEV FLGLGSLILDPNIFKHPFNPDRFLADGRFKEAFPLFSGLKRVCLGEGIAKAEI FLFFTLILQAFSLSPCPDPTLSLKPTVSLFNIPPAFLQVRPTDLHSTTQTR"  
177..1493  
/note="p450; Region: Cytochrome P450. Cytochrome P450s are involved in the oxidative degradation of various compounds. Particularly well known for their role in the degradation of environmental toxins and mutagens. Structure is mostly alpha, and binds a heme cofactor"  
/db\_xref="CDD:pfam00067"

ORIGIN

Alignment Scores:

Pred. No.: 5.3e-201 Length: 2619  
Score: 2615.00 Matches: 504  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 9 Gaps: 0

US-10-669-693-2 (1-504) x BC033691 (1-2619)

QY 1 MetGluAlaThrGlyThrTrpAlaLeuLeuAlaLeuLeuLeuLeuLeuLeuThr 20  
Db 18 ATGAGCGGACCGGACCTGGCGCTGCTGGCGCTGGCGCTGGCGCTGGCGCTGACG 77  
QY 21 LeuAlaLeuSerGlyThrArgAlaArgGlyHisLeuProGlyProThrProLeuPro 40  
Db 78 CTGGCGCTGTCGGGACCGAGGCGCCGAGGCCACCTGCCCGCGGCCACCGCGCTACCA 137  
QY 41 LeuLeuGlyAsnLeuLeuGlnLeuArgProGlyAlaLeuTyrsGlyLeuMetArgLeu 60  
Db 138 CTGCTGGAAACCTCTCTGAGCTACGGCGCGCGCTGTATTTCAGGGCTCATGCGGCTG 197  
QY 61 SerLysLysTrpGlyProValPheThrIleTyrsLeuGlyProTrpArgProValValVal 80  
Db 198 AGTAAGAAGTACCGACCGGTTCACCATCTACTGGGACCCCTGGCGGCTGTGGTGTGTC 257  
QY 81 LeuValGlyGlnGluAlaValArgGluAlaLeuGlyGlyGlnAlaGluGluPheSerGly 100  
Db 258 CTGTTGGGAGGAGGCTGTGCGGAGGCCCTGGAGGCTGAGGCTGAGGAGTTTCAGCGGC 317  
QY 101 ArgGlyThrValAlaMetLeuGluGlyHisGlyValPhePheSerAsn 120  
Db 318 CGGGAAACCGTAGCATGCTGGAAGGACTTTTGTATGCCATGGGGTTTCTTCTCCAC 377  
QY 121 GlyClnuArgTrpArgGlnLeuArgGlyPheThrMetLeuAlaLeuArgAspLeuGlyMet 140  
Db 378 GGGAGCGGTGGAGGACCTGAGAAAGTTTACCATGTTGCTCTCGCGGACCTGGGCATG 437  
QY 141 GlyLysArgGluGlyGluGluLeuIleGlnAlaGluAlaArgCysLeuValGluThrPhe 160  
Db 438 GGGAGCGAGAAAGCGAGGAGCTGATCCAGCGGAGGCCCGCTGTCTGTGGAGACATTC 497  
QY 161 GlnGlyThrGluGlyArgPropPheAspProSerLeuLeuAlaGlnAlaThrSerAsn 180



Db 498 CAGGGGACAGAGGACGCGCATTCGATCCCTCGCTGCTGGCCAGGCGCACCTCCAC 557  
 Qy 181 ValValCysSerLeuLeuPheGlyLeuArgPheSerTyrGluAspLysGluPheGlnAla 200  
 Db 558 GTAGTCTGCTCCCTCTTTGGCTCGCGCTTCTCTATGAGGATAAGGAGTTCAGGCGC 617  
 Qy 201 ValValArgAlaAlaGlyGlyThrLeuLeuGlyValSerSerGlnGlyGlyGlnThrTyr 220  
 Db 618 GTGGTCCGGGACGTGGTGTACCTCGCTGGAGTTCAGTCCAGGGGGTTCAGACCTAC 677  
 Qy 221 GluMetPheSerTrpPheLeuArgProLeuProGlyProHisLysGlnLeuLeuHisHis 240  
 Db 678 GAGATGTTCTCTGCTGCTCGCGCCCTGCGAGCCGCCCAAGCAGCTCTCCACAC 737  
 Qy 241 ValSerThrLeuAlaAlaPheThrValArgGlnValGlnHisGlnGlyAsnLeuAsp 260  
 Db 738 GTACGACCTTGGCTGCTTCCAGTCCGGCAGGTGCAGCAGCACCCAGGGGAACCTGGAT 797  
 Qy 261 AlaSerGlyProAlaArgAspLeuValAspAlaPheLeuLeuLysMetAlaGlnGluGlu 280  
 Db 798 GCTTCGGGGCCCGCAGCGTACCTTGTGATGCTTCTGCTGAAGATGGCACAGAGGAA 857  
 Qy 281 GlnAsnProGlyThrGluPheThrAsnLysAsnMetLeuMetThrValIleTyrLeuLeu 300  
 Db 858 CAAGACCCAGGCAAGAAATTCACCAAGAAACATGCTGATGACAGTCAATTTATTTGCTG 917  
 Qy 301 PheAlaGlyThrMetThrValSerThrThrValGlyTyrThrLeuLeuLeuMetLys 320  
 Db 918 TTTGCTGGGAGATGACGGTCAGCACCGGTGCGGTATACCCCTCTGCTCTGATGAAA 977  
 Qy 321 TyrProHisValGlnLysTrpValArgGluGluLeuAsnArgGluLeuGlyValGlyGln 340  
 Db 978 TAGCCTCATGTCACAAAGTGGTACGTGAGGAGCTGAATCGGAGCTGGGGCTGGCCAG 1037  
 Qy 341 AlaProSerLeuGlyAspArgThrArgLeuProTyrThrAspAlaValLeuHisGluAla 360  
 Db 1038 GCACCAAGCCTTAGGGACCGTACCGCTCTCTTACACCGAGCGGTTCGTGATGAGCGC 1097  
 Qy 361 GlnArgLeuLeuAlaLeuValProMetGlyLeuProArgThrLeuMetArgThrArg 380  
 Db 1098 CAGCGGCTGCTGGCGGTGGTGGCCATGGGAATACCCCGCACCTCATGCGGACCAACCCG 1157  
 Qy 381 PheArgGlyTyrThrLeuProGlnGlyThrGluValPheProLeuLeuGlySerIleLeu 400  
 Db 1158 TTCGAGGGGTACACCTGCCCCAGGGACCGAGGTCTTCCCTCTCTCTGCTCATCTCTG 1217  
 Qy 401 HisAspProAsnIlePheLysHisProGluGluPheAsnProAspArgPheLeuAspAla 420  
 Db 1218 CATGACCCCAACATCTTCAAGCACCCAGAGAGTTCACCCAGACCGTTTCTTGGATGCA 1277  
 Qy 421 AspGlyArgPheArgLysHisGluAlaPheLeuProPheSerLeuGlyLysArgValCys 440  
 Db 1278 GATGACGGTTCAGGAAGCATGAGCGGTCTCTGCTCTCTTCTTCTTCTTCTTCTGTCG 1337  
 Qy 441 LeuGlyGluGlyLeuAlaLysAlaGluLeuPhePheThrThrIleLeuGlnAla 460  
 Db 1338 CTTGGAGAGCGCTGCAAAAGCGGAGCTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1397  
 Qy 461 PheSerLeuGluSerProCysProAspThrLeuSerLeuLysProThrValSerGly 480  
 Db 1398 TTTCTCTCTGAGAGCGCGTGGCCCGCGGACACCTGAGCCTCAAGCCCAACCTTCTTCTTCT 1457  
 Qy 481 LeuPheAsnIleProProAlaPheGlnLeuGlnValArgProThrAspLeuHisSerThr 500  
 Db 1458 CTTTTCACATTTCCCCCAGCTTTCAGCTGCAAGTCCGTCCCACTGACCTTCACTCCACC 1517  
 Qy 501 ThrGlnThrArg 504  
 Db 1518 ACGCAGACCCAGA 1529  
 RESULT 6  
 AX327331

LOCUS AX327331 2620 bp DNA linear PAT 07-JAN-2002  
 DEFINITION Sequence 16 from Patent WO0179468.  
 ACCESSION AX327331  
 VERSION AX327331.1 GI:18097877  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 Policky,J.L., Hafalia,A., Burford,N., Ring,H.Z., Lal,P.,  
 Tribouley,C.M., Yao,M.G., Yue,H., Tang,Y.I., Patterson,C., Das,D.,  
 Sanjanwala,M.S., Gandhi,A.K., Reddy,R., Khan,F.A., Baughn,M.R.,  
 Rankumar,J., Griffin,J.A. and Au-Young,J.  
 Drug metabolizing enzymes  
 Patent: WO 0179468-A 16 25-OCT-2001;  
 Incyte Genomics, Inc. (US)  
 FEATURES  
 Location/Qualifiers  
 source 1..2620  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"  
 /note="Incyte ID No: 1272843CB1"  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 5,3e-201 Length: 2620  
 Score: 2615.00 Matches: 504  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 6 Gaps: 0  
 US-10-669-693-2 (1-504) x AX327331 (1-2620)  
 Qy 1 MetGluAlaThrGlyThrTrpAlaLeuLeuAlaLeuAlaLeuLeuLeuLeuLeuThr 20  
 Db 41 ATGGAGCGACCGGACCGGCGCTGCTGCTGGCGTGGCGTCTCTCTGCTGCTGCTGCTG 100  
 Qy 21 LeuAlaLeuSerGlyThrArgAlaArgGlyHisLeuProGlyProGlyProThrProLeuPro 40  
 Db 101 CTGGCGCTGTCCGGACCGAGGCGCCGAGGCCACTGCCCCCGGGCCCGCGCTACCA 160  
 Qy 41 LeuLeuGlyAsnLeuLeuGlnLeuArgProGlyValAlaLeuTyrSerGlyLeuMetArgLeu 60  
 Db 161 CTGCTGGGAAACCTCTCTGAGCTACGCGCGCGGCGCTGTATTTCAGGGCTCATCGCGCTG 220  
 Qy 61 SerLysLysTyrGlyProValPheThrIleTyrLeuGlyProTrpArgProValValVal 80  
 Db 221 AGTAAGAGTACGAGACCGGTGTTACCATCTACTGCGACCTTGGCGGCTGTGGTGGTGC 280  
 Qy 81 LeuValGlyGlnGluAlaValArgGluAlaLeuGlyGlyGlnAlaGluGluPheSerGly 100  
 Db 281 CTGGTTGGGACGAGGCTGTGGGAGGCGCTTGGAGGTCAGGCTGAGGAGTTCAGCGGC 340  
 Qy 101 ArgGlyThrValAlaMetLeuGluGlyThrPheAspGlyHisGlyValPhePheSerAsn 120  
 Db 341 CGGGGAACCGTAGCATGCTGGAAGGAGCTTTTGTATGGCCATGGGGTCTTCTTCTCAAC 400  
 Qy 121 GlyGluArgTrpArgGlnLeuArgLysPheThrMetLeuAlaLeuArgAspLeuGlyMet 140  
 Db 401 GGGGACGCGTGGAGGACGCTGAGGAAGTTTACCATGCTTGTCTGCGGACCTTGGGCGATG 460  
 Qy 141 GlyLysArgGluGlyGluLeuIleGlnAlaGluAlaArgCysLeuValGluThrPhe 160  
 Db 461 GGGAGCGAGAGAGCGAGGCTGATCCAGGCGAGGCGCGGTGTCTGTGGAGACATTTC 520  
 Qy 161 GlnGlyThrGluGlyArgProPheAspProSerLeuLeuAlaGlnAlaThrSerAsn 180  
 Db 521 CAGGGGACAGAGGAGCGCCATTCATCCCTCCCTGCTGCTGGCCCGAGCCACCTCCCAAC 580  
 Qy 181 ValValCysSerLeuLeuPheGlyLeuArgPheSerTyrGluAspLysGluPheGlnAla 200

581 GTAGTGTGCTCCCTCTCTTTGGGCTCGCTTCTCTATGAGGATAAGGAGTTCACGACC 640  
201 ValValArgAlaAlaGlyThrLeuLeuGlyValSerSerGlnGlyGlyGlnThrTyr 220  
641 GTGGTCCGGGAGCTGGTGGTACCTCTGTGGAGTCAGTCCAGGGGGTTCAGACCTAC 700  
221 GluMetPheSerTrpPheLeuArgProLeuProGlyProHisLysGlnLeuLeuHis 240  
701 GAGATGTTCTCTGCTGCTGGGCTCTGCGGCTCTGCGGCTCTGCGGCTCTGCGGCTCT 760  
241 ValSerThrLeuAlaAlaPheThrValArgGlnValGlnHisGlnGlnValAsnLeuAsp 260  
761 GTCAGACCTTGGCTGGCTTTCACATGTCGGCAGGTGCAGACACACAGGGAGACCTGGAT 820  
261 AlaSerGlyProAlaArgAspLeuValAlaPheLeuLeuLysMetAlaGlnGluGlu 280  
821 GCTTCGGGCTCCGACGTCGACCTTGTGATGCTTCTGCTGCAAGATGTCACAGGAGAA 880  
281 GlnAsnProGlyThrGluPheThrAsnLysAsnMetLeuMetThrValIleTyrLeuLeu 300  
881 CAAACCCAGGACACGAAATTCACCAAGAACATGCTGATGACAGTCAATTTATTGTG 940  
301 PheAlaGlyThrMetThrValSerThrValGlyTyrThrLeuLeuLeuMetLys 320  
941 TTTGCTGGGACGATGACGTCAGCACCCGCTCGGCTATACCTCTCTCTCTGATGAA 1000  
321 TyrProHisValGlnLysTrpValArgGluGluLeuAsnArgGluLeuGlyAlaGln 340  
1001 TACCCTCATGTCACAAAGTGGTACGTGAGGAGTGAATCGGAGCTGGGGCTGCCAG 1060  
341 AlaProSerLeuGlyAspArgThrArgLeuProTyrThrAspAlaValLeuHisGluAla 360  
1061 GCACCAAGCTTAGGGACCGTACCCGCTCCCTTACACGACGCGTCTCTGATGAGGG 1120  
361 GlnArgLeuLeuAlaLeuValProMetGlyIleProArgThrLeuMetArgThrThrArg 380  
1121 CAGCGCTGCTGGCGCTGGTGGCCATGGGAAATACCCGACCCCTCATGGGACCAACCCG 1180  
381 PheArgGlyTyrThrLeuProGlnGlyThrGluValPheProLeuLeuGlySerIleLeu 400  
1181 TTCGAGGTATACCTGCTGCCAGGACGAGGACGAGGCTTCTCCCTCTCTTGGCTCATCTG 1240  
401 HisAspProAsnIlePheLysHisProGluGluPheAsnProAspArgPheLeuAspAla 420  
1241 CATGACCCCAACATCTTCAAGACCCAGAGAGTTCAACCCAGACCGTTTCTCGGATGCA 1300  
421 AspGlyArgPheArgLysHisGluAlaPheLeuProPheSerLeuGlyLysArgValCys 440  
1301 GATGGACGTTTCAGGAAGCATGAGGCTTCTGCTCTCTCTTCTTCCATCCATCTACAAGC 1360  
441 LeuGlyGluGlyLeuAlaLysAlaGluLeuPhePheThrThrIleLeuGlnAla 460  
1361 CTTGGAGAGGCTGGCAAGGAGGAGTCTTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCT 1420  
461 PheSerLeuGluSerProCysProAspThrLeuSerLeuLysProThrValSerGly 480  
1421 TTTCTCTGAGAGCCGTCGCGCGCGGACACCTGAGCCTCAAGCCACCGTCAGTGGC 1480  
481 LeuPheAsnIleProProAlaPheGlnLeuGlnValArgProThrAspLeuHisSerThr 500  
1481 CTTTTCACATTTCCCCAGCTTCCAGCTGCAAGTCCGTCCTCCACTGACCTTCACTCCACC 1540  
501 ThrGlnThrArg 504  
1541 ACGCAGACCAGA 1552

RESULT 7  
AX463930  
LOCUS  
DEFINITION Sequence 63 from Patent WO0140466.  
ACCESSION AX463930  
VERSION AX463930.1  
KEYWORDS GI:21898982

SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE  
AUTHORS Baker, K.P., Beresini, M., Deforge, L., Desnoyers, L., Filvaroff, E.,  
Gao, W.Q., Gerritsen, M.E., Goddard, A., Godowski, P.J., Gurney, A.L.,  
Sherwood, S., Smith, V., Stewart, T.A., Tumas, D., Watanabe, C.K.,  
Wood, W.L. and Zhang, Z.  
TITLE Secreted and transmembrane polypeptides and nucleic acids encoding  
same  
JOURNAL Patent: WO 0140466-A 63 07-JUN-2001;  
Genentech Inc. (US)  
FEATURES  
source 1. 2623  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Alignment Scores: 5.31e-201 Length: 2623  
Pred. No.: 2615.00 Matches: 504  
Score: 100.00% Conservative: 0  
Percent Similarity: 100.00% Mismatches: 0  
Best Local Similarity: 100.00% Indels: 0  
Query Match: 100.00% Gaps: 0  
DB: 6  
US-10-669-693-2 (1-504) x AX463930 (1-2623)  
QY 1 MetGluAlaThrGlyThrTrpAlaLeuLeuLeuAlaLeuLeuLeuLeuLeuLeuLeuThr 20  
Db 41 ATGGAGGCGACCGCACCTGGCGCTGCTGCTGGCGCTGCTGCTGCTGCTGCTGCTGCTG 100  
QY 21 LeuAlaLeuSerGlyThrArgAlaArgGlyHisLeuProGlyProGlyProThrProLeuPro 40  
Db 101 CTGGCGCTCTCGCGGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 160  
QY 41 LeuLeuGlyAsnLeuLeuGlnLeuArgProGlyAlaLeuTyrSerGlyLeuMetArgLeu 60  
Db 161 CTGCTGGGAAACCTCTCGAGTACGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGG 220  
QY 61 SerLysLysTyrGlyProValPheThrIleTyrLeuGlyProThrArgProValValVal 80  
Db 221 AGTAAGAAGTACGAGCGCGGCTTACCATCTACCTGGGACCCCTGGCGGCGGCTGCTG 280  
QY 81 LeuValGlyGlnGluAlaValArgGluAlaLeuGlyGlyGlnAlaGluGluPheSerGly 100  
Db 281 CTGGTTGGGCGAGGAGCTGTGGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 340  
QY 101 ArgGlyThrValAlaMetLeuGluGlyThrPheAspGlyHisGlyValPhePheSerAsn 120  
Db 341 CGGGGAACCTGACGATGCTGGAGGAGACTTTTGTATGGCCATGGGCTTTTCTTCTTCT 400  
QY 121 GlyGluArgTrpArgGlnLeuArgLysPheThrMetLeuAlaLeuArgAspLeuGlyMet 140  
Db 401 GGGGAGCGGTGGAGGAGCTGAGGAAGTTTACCATGCTTCTGCTGCGGAGCTTGGGCG 460  
QY 141 GlyLysArgGluGlyGluGluLeuIleGlnAlaGluAlaArgCysLeuValGluThrPhe 160  
Db 461 GGGGAAGCGAAGCGGAGGAGCTGATCCAGCGGAGGCGGCGGCGGCGGCGGCGGCGG 520  
QY 161 GlnGlyThrGluGlyArgProPheAspProSerLeuLeuLeuAlaGlnAlaThrSerAsn 180  
Db 521 CAGGGACAGAGGACCGCCCATTCGATCCCTCTCTGCTGCGGCGGCGGCGGCGGCGG 580  
QY 181 ValValCysSerLeuLeuPheGlyLeuArgPheSerTyrGluAspLysGluPheGlnAla 200  
Db 581 GTAGTGTCTCTCTCTTGGCGCTTCTCTTATGAGGATAAGGAGTTCAGGCGC 640  
QY 201 ValValArgAlaAlaGlyGlyThrLeuLeuGlyValSerSerGlnGlyGlyGlnThrTyr 220  
Db 641 GTGGTCCGGGCGAGCTGTGTGTATCCCTGCTGGGAGTCAGCTCCCGGCGGCGGCGG 700

Qy	221	GluMetPheSerTrpPheLeuArgProLeuProGlyProHisLysGlnLeuLeuHis	240
Db	701	GAGATGTTCTCTGGTTCCTGGGCCCTTCAGGCCCCCAAGCAGCTCTCTCCACCAC	760
Qy	241	ValSerThrLeuAlaAlaPheThrValArgGlnValGlnGlnHisGlnGlyAsnLeuAsp	260
Db	761	GTCAGCACCTTGGTGCTTCACAGTCCGGCAGGTGCAGCAGCACAGGGGAACCTGGAT	820
Qy	261	AlaSerGlyProAlaArgAspLeuValAspAlaPheLeuLeuLysMetAlaGlnGluGln	280
Db	821	GC TTCGGGGCCCGCAGCTGACCTTGTGATGCCCTTCCTGCTGAAGATGGCAGAGGAA	880
Qy	281	GlnAsnProGlyThrGluPheThrAsnLysAsnMetLeuMetThrValIleTyrLeuLeu	300
Db	881	CAAAACCCAGGCACAGAAATTCACCAACAGAACATGCTGATGACAGTCAATTAATTGCTG	940
Qy	301	PheAlaGlyThrMetThrValSerThrThrValGlyTyrThrLeuLeuLeuLeuMetLys	320
Db	941	TTTGTCTGGGACGATACCGTCAGCACACCGTCGGCTATACCTTCCTGCTCTGTGATGAA	1000
Qy	321	TyrProHisValGlnLysTyrValArgGluLeuLeuAsnArgGluLeuGlyAlaGlyGln	340
Db	1001	TACCCCTATGTCAAAAGTGGTACGTGAGGAGCTGAATCGGAGCTGGGGCTGGCCAG	1060
Qy	341	AlaProSerLeuGlyAspArgThrArgLeuProTyrThrAspAlaValLeuHisGluAla	360
Db	1061	GCACCAAGCCTAGGGGACCGTACCGGCTCCCTTACACCGACGCGTTCCTGCAAGGCG	1120
Qy	361	GlnArgLeuLeuAlaLeuValProMetGlyIleProArgThrLeuMetArgThrThrArg	380
Db	1121	CAGCGGCTGTGGCGTGTGGTCCCATGGGAATACCCCGCACCTCATGCGGACACCCGC	1180
Qy	381	PheArgGlyTyrThrLeuProGlnGlyThrGluValPheProLeuLeuGlySerIleLeu	400
Db	1181	TTCCGAGGGGTACACCTGCCCGCAGGGCAGGAGGTCTTCGCCCTCTTGCTCCATCTTG	1240
Qy	401	HisAspProAsnIlePheLysHisProGluGluPheAsnProAspArgPheLeuAspAla	420
Db	1241	CATGACCCCAACATCTTCAGCACCCAGAAAGATTCAAACCCAGACCGTTTCCTGGATGCA	1300
Qy	421	AspGlyArgPheArgLysHisGluAlaPheLeuProPheSerLeuGlyLysArgValCys	440
Db	1301	GATGACCGTTCAGGAAGCATGAGGCGTTCCTGCCCTTCTCCTTAGGAAGGCGTGCTCG	1360
Qy	441	LeuGlyGluGlyLeuAlaLysAlaGluLeuPhePheThrThrIleLeuGlnAla	460
Db	1361	CTTGAGAGGGCTGGCAAAAGCGGAGCTCTTCTCTTCTTCCACCAACCTCTACAAGCC	1420
Qy	461	PheSerLeuGluSerProCysProProAspThrLeuSerLeuLysProThrValSerGly	480
Db	1421	TTCTCCCTGGAGGCCCGTGCCCGCGGACACCTTGAGCTTCAAGGCCACCGTCAGTGCC	1480
Qy	481	LeuPheAsnIleProProAlaPheGlnLeuGlnValAtgProThrAsnLeuHisSerThr	500
Db	1481	CTTTTCAACATTCCTCCAGCCTTCACAGTGCAGGTCGTCGAGTCCGTCACCTTCACTCCAC	1540
Qy	501	ThrGlnThrArg	504
Db	1541	ACGCAGACCAAGA	1552

RESULT 8  
 AY358603  
 LOCUS  
 DEFINITION  
 AY358603  
 Homo sapiens clone DNA76395 Cytochrome p450 (UQ0891) mRNA, complete  
 cds.  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 AY358603  
 Homo sapiens (human)  
 FLI\_CDNA.  
 GI:37182327  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 2626)  
Clark,H.F., Gurney,A.L., Abaya,E., Baker,K., Baldwin,D., Brush,J.,  
Chen,J., Chow,B., Chui,C., Crowley,C., Currell,B., Deuel,B.,  
Dowd,P., Eaton,D., Foster,J., Grimaldi,C., Gu,Q., Hass,P.E.,  
Heldens,S., Huang,A., Kim,H.S., Klimowski,L., Jin,Y., Johnson,S.,  
Lee,J., Lewis,L., Liao,D., Mark,M., Robbie,E., Sanchez,C.,  
Schoenfeld,J., Seshagiri,S., Simmons,L., Singh,J., Smith,V.,  
Stinson,J., Vagts,A., Vanden,R., Watanabe,C., Wileand,D., Woods,K.,  
Xie,M.H., Yansura,D., Yi,S., Yu,G., Yuan,J., Zhang,M., Zhang,Z.,  
Goddard,A., Wood,W.I. and Godowski,P.  
The Secreted Protein Discovery Initiative (SPDI), a Large-Scale  
Effort to Identify Novel Human Secreted and Transmembrane Proteins:  
A Bioinformatics Assessment  
Genome Res. 13 (10), 2265-2270 (2003)  
2 (bases 1 to 2626)  
Clark,H.F.  
Direct Submission  
Submitted (01-AUG-2003) Department of Bioinformatics, Genentech,  
Inc., 1 DNA Way, South San Francisco, CA 94080, USA.  
Location/Qualifiers  
1. .2626  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="DNA76395"  
1. .2626  
/locus\_tag="UNQ891"  
44. .1558  
/locus\_tag="UNQ891"  
/note="PRO1906"  
/codon\_start=1  
/product="Cytochrome p450"  
/protein\_id="AAQ88966.1"  
/db\_xref="GI:37182328"  
/translation="MEATGTWALLIALALLLLTALSQRHGLPGPTPLLLGN  
LQLRPGALYGLMRLLSKYGVFTYILGPWRPVVLVGQAVREALGGQAEFSRG  
TVAMLEGDFDGHVFFSNGERNQLRKFTRLALRDLMGMRCEGELLQAEARCLVTF  
QCTEGRFDPSGLLLAQATSNVVCSLILFGLRFSVDEKFEQFVAPRAGTILGVSSQGGQ  
TVEMFSWFLRPLPGPKQLLHVSTLAATFVQVQHQNLDSAGPARDLVDAFLIKM  
AQEONQCFTEFTNKNMLMTVYILFAGTMTVSTVGYTLLLLMKTPHYQKVRBELRN  
ELGACQASLGDRTRLPYDAVLLHEAQRIALLVPMGIPTRLMTRTFRFGYITLPQGTVE  
FLLGLSLHDHNPFIKHPPEFNDRLDADGRFRKHEAFLPFLSLGKRVCLGEGGLAKEL  
FLFTTIIQLQAFSLSPCPDITLSLKPXTVSGLENIIPPAFOI OVRPTDLSHTTTPP"

## ORIGIN

Alignment Scores:		
Pred. No.:	5.32e-201	Length: 2626
Score:	2615.00	Matches: 504
Percent Similarity:	100.00%	Conservative: 0
Best Local Similarity:	100.00%	Mismatches: 0
Query Match:	100.00%	Indels: 0
DB:	9	Gaps: 0

US-10-669-693-2 (1-504) X AY358603 (1-2626)

[illegible]

Db 1364 CTTGGAGAGGCGCTTGTCAAAGCGGAGCTCTTCCTCTTACCACCATCCTACAAGCC 1423

Qy 461 PheSerLeuGluSerProCysProAspThrLeuSerLeuLysProThrValSerGly 480

Db 1424 TTCTCCCTTGAGAGCGCTGCCGCGGACACACCTGAGCTCAAGCCACCGTCAGTGGC 1483

Qy 481 LeuPheAsnIleProProAlaPheGlnLeuGlnValArgProThrAspLeuHisSerThr 500

Db 1484 CTTTTAACATTCCTCCCAGGCCTTCAGCTGAAGTCCGCTCCACTGACCTTCACCTCAC 1543

Qy 501 ThrGlnThrArg 504

Db 1544 ACGCAGACCAAGA 1555

RESULT 9  
AX552230 2604 bp DNA linear PAT 27-NOV-2002  
LOCUS Sequence 139 from Patent WO0162927.  
DEFINITION AX552230  
ACCESSION AX552230 GI:25896468  
VERSION  
KEYWORDS Homo sapiens (human)  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS Banville,S.C., Greenwalt,L.B., Lincoln,S.E., Stockdreher,T.K., Amshvey,S., Chang,S.C., Chen,W., D'Sa,S.A., Dam,T.C., Liu,T.F., Rosen,B.H., Russo,F.D., Spiro,P.A., Bradley,D.L., Chen,A., Cohen,H.J., Daffo,A., Daniels,S.E., Dufour,G.E., Flores,V., Fong,W.T., Hodgson,D.M., Jackson,S., Jones,A.L., Panzer,S., Roseberry,A.M., Shah,P., Wright,R.J., Yap,P.E., Yu,J.Y., Bratcher,S.R., Chalup,M.S., Dahl,C.R. and Hillman,J.L.  
TITLE Polypeptides and corresponding polynucleotides for diagnostics and therapeutics  
JOURNAL Patent: WO 0162927-A 139 30-AUG-2001;  
INCYTE Genomics, Inc. (US)  
FEATURES  
Location/Qualifiers  
1..2604  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
/note="Incute ID No: LG:337818.2:2000FEB18"

ORIGIN  
Alignment Scores:  
Pred. No.: 1,33e-200 Length: 2604  
Score: 2610.00 Matches: 503  
Percent Similarity: 99.80% Conservative: 0  
Best Local Similarity: 99.80% Mismatches: 1  
Query Match: 99.81% Indels: 0  
DB: Gaps: 6

US-10-669-693-2 (1-504) x AX552230 (1-2604)

Qy 1 MetGluAlaThrGlyThrTrpAlaLeuLeuLeuAlaLeuLeuLeuLeuLeuLeuLeuThr 20

Db 40 ATGAGCGCAGCCACCTCTGTCAGCTACGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTG 99

Qy 21 LeuAlaLeuSerGlyThrArgAlaArgGlyHisLeuProProGlyProThrProLeuPro 40

Db 100 CTGGCGCTGTCTGGGACACGAGGCGCCAGGCGACCTGCCCCCGGCGCGCTACCA 159

Qy 41 LeuLeuGlyAsnLeuLeuGlnLeuArgProGlyAlaLeuTyrsSerGlyLeuMetArgLeu 60

Db 160 CTGCTGGAAACCTCTGTCAGCTACGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTG 219

Qy 61 SerLysLysTyrGlyProValPheThrIleTyreGlyProThrArgProValValVal 80

Db 220 AGTAAGAAGTACGAGCGGTGTTACCATCTACTGGGACCTCTGGGCGCTCTGTTGGTC 279

Qy 81 LeuValGlyGlnGluAlaValArgGluAlaLeuGlyGlyGlnAlaGluPhSerGly 100

Db 280 CTGGTGGGACGAGGCTGTGCGGAGGCCCCCTGGGAGGTTCAGGCTTCAGCGGC 339  
 QY 101 ArgGlyThrValAlaMetLeuGluGlyThrPheAspGlyHisGlyValPhePheSerAsn 120  
 Db 340 CGGGACCGTAGCGATGCTGGAGGAGCTTTTGATGGCCATGGGGTTTCTTCTCCAC 399  
 QY 121 GlyGluArgTyrArgGlnLeuArgLysPheThrMetLeuAlaLeuArgAspLeuGlyMet 140  
 Db 400 GGGGAGCGGTGAGGACGTCAGGAAGTTTACCATGCTTGTCTCTCGGGACCTGGGCATG 459  
 QY 141 GlyLysArgGluGlyGluGluLeuLeuAlaGluAlaArgCysLeuValGluThrPhe 160  
 Db 460 GGGAGCGGAGAGCGAGGAGCTGATCCAGGCGGAGGCCGCTGTCTGTGGAGACATTC 519  
 QY 161 GlnGlyThrGluGlyArgProPheAspProSerLeuLeuLeuAlaGlnAlaThrSerAsn 180  
 Db 520 CAGGGACAGAGAGCGCCATTCGATCCCTCCCTGCTGTGGCCAGGCCACCTCCAC 579  
 QY 181 ValValCysSerLeuLeuPheGlyLeuArgPheSerTyrGluAspLysGluPheGlnAla 200  
 Db 580 GTAGTCTGCTCCCTCTTGGGCTCGCTTCTCTATGAGGATAGGAGTTCCAGGCC 639  
 QY 201 ValValArgAlaAlaGlyGlyThrLeuLeuGlyValSerSerGlnGlyGlyGlnThrTyr 220  
 Db 640 GTGTCCGGGAGCTGTGTGTACCTGCTGGGAGTTCAGTCCAGGGGGGTTCAGACCTAC 699  
 QY 221 GluMetPheSerTrpPheLeuArgProLeuProGlyProHisLysGlnLeuLeuHis 240  
 Db 700 GAGATGTTCTCTGTGTCCTGGGCCCTGCGAGGCCCCACAGCAGCTCTCCACAC 759  
 QY 241 ValSerThrLeuAlaAlaPheThrValArgGlnValGlnGlnHisGlnGlyAsnLeuAsp 260  
 Db 760 GTCAGCACCTTGGCTGCTTCCACAGTCCGCGAGTGCAGCAGCACACAGGGAACCTGGAT 819  
 QY 261 AlaSerGlyProAlaArgAspLeuValAspAlaPheLeuLeuLysMetAlaGlnGlu 280  
 Db 820 GCTTCGGCCCCGACGTGATGCTGTGATGCTTCTGCTGAGATGGCACAGGAGAA 879  
 QY 281 GlnAsnProGlyThrGluPheThrAsnLysAsnMetLeuMetThrValIleTyrLeuLeu 300  
 Db 880 CAAACCCAGGCACAGATTCACCAACAGACATGCTGATGACAGTCATTTATTGTGCTG 939  
 QY 301 PheAlaGlyThrMetThrValSerThrThrValGlyTyrThrLeuLeuLeuMetLys 320  
 Db 940 TTTGCTGGGACGATGCGGTGACGACCCAGTTCGCTATACCTCTCTGCTCTGATGAAA 999  
 QY 321 TyrProHisValGlnLysTrpValArgGluGluLeuAsnArgGluLeuGlyAlaGlyGln 340  
 Db 1000 TACCCCTCATGTCAAAGTGGGTACGTGAGGAGCTGAAATCGGAGCTGGGGGTGGCCAG 1059  
 QY 341 AlaProSerLeuGlyAspArgThrArgLeuProTyrThrAspAlaValLeuHisGluAla 360  
 Db 1060 GCACCAAGCCTAGGGACCGTACCCGCTCCCTTACCGACGCGGTTCGTGATGAGCGC 1119  
 QY 361 GlnArgLeuLeuAlaLeuValProMetGlyLeuProArgThrLeuMetArgThrThrArg 380  
 Db 1120 CAGCGCTGCTGCGGTGCTGCCATGGAATACCCCGCACCTCTATGCGGACCCCGC 1179  
 QY 381 PheArgGlyTyrThrLeuProGlnGlyThrGluValPheProLeuLeuGlySerIleLeu 400  
 Db 1180 TTCCGGGGGTACACCTGCGCCAGGGCACGGAGGTCTTCCGCTCTGCTGGCTCCATCTG 1239  
 QY 401 HisAspProAsnIlePheLysHisProGluGluPheAsnProAspArgPheLeuAspAla 420  
 Db 1240 CATGACCCCAACATCTTCAGCACCCAGAGAGATTCAACCCAGACCGTTTCTGTGATGCA 1299  
 QY 421 AspGlyArgPheArgLysHisGluAlaPheLeuProPheSerLeuGlyLysArgValCys 440  
 Db 1300 CATGACGCTTCAGGAAGCATCAGGCGTTCCTGCGCTTCTCTTAGGAGAGCGTGTCTGC 1359  
 QY 441 LeuGlyGluGlyLeuAlaLysAlaGluLeuPhePheThrThrIleLeuGlnAla 460  
 Db 1360 CTTGGAGAGGGCTGGCAAAAGCGAGGTGCTTCTCTTCTTCAACACCATCTTACAGCC 1419

QY 461 PheSerLeuGluSerProCysProAspThrLeuSerLeuLysProThrValSerGly 480  
 Db 1420 TTTCTCCCTGGAGAGCCGTCGCCGCGGACACCTTAGCCTCAAGCCACACCTCAGTGC 1479  
 QY 481 LeuPheAsnIleProProAlaPheGlnLeuGlnValArgProThrAspLeuHisSerThr 500  
 Db 1480 CTTTTCACATTTCCCCAGCCTTCCAGCTGCAAGTCCGTCCTCCACTGACCTTCACTCCAC 1539  
 QY 501 ThrGlnThrArg 504  
 Db 1540 ACGCAGACCAGA 1551  
 RESULT 10  
 AX876954 2513 bp DNA linear PAT 17-DEC-2003  
 LOCUS  
 DEFINITION Sequence 11859 from Patent EP1074617.  
 ACCESSION AX876954  
 VERSION AX876954.1 GI:40031690  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1  
 AUTHORS Ota, T., Isogai, T., Nishikawa, T., Hayashi, K., Saito, K., Yamamoto, J.,  
 Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K. and Otsuki, T.  
 TITLE Primers for synthesizing full-length cDNA and their use  
 JOURNAL Patent: Ep 1074617-A 11859 07-FEB-2001;  
 Research Association for Biotechnology (JP)  
 FEATURES  
 source Location/Qualifiers  
 1. 2513  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"  
 56..1750  
 /notes="unnamed protein product"  
 /codon\_start=1  
 /protein\_id="CAE89506.1"  
 /db\_xref="GI:40031691"  
 /translation="MEATGTWALLLALLLLTALSGTRARGHLLPGPTPLPLGN  
 LLIQRRGALYSGLMRUSKIGPVFTLLYLPWRPVVVLGVQEAVERALGGQAEFSRG  
 TVAMLEGTDFGHVFFSNGERWQLRFTMLALRDLGMKREGEELTQAEARCLVETP  
 QGTGEPFPDPLLAQATSNVVCSLIFGLRFVDEKFOAVRAAGGTLGLVSSQSGQ  
 TVEMSFWRPLPGPHKQLLHVSTLAAFTVRQVQOHQNLDAAGPARDLVADPLAKM  
 AOEONPQTEFTNKMMLTVIYLLFAGTWTSTVTCGYTLLLLMKYPHVQKWRBELNR  
 ELGAGQAPSLGDRTRLPYDVLHQAORLLALVPMGIPTLMTTRFRGTYLLPQGTVEV  
 FPLLSILHDPNI FKHPSEFNDFLDADGRFRKHEAFLLFSLGKRVCLGEGIAKEL  
 FLFTTLLQAFSLSPCPDPTLSLKPFTVSGLFNIPPASTVGMDRVNVSVYTAGSHY  
 TPAVFRSLSHGPHALHTAAKMHNRTPIHNYKGHKATAGLAFHHRHKYSPSAITST"  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 6, 2e-194 Length: 2513  
 Score: 2527.00 Matches: 487  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 96.63% Indels: 0  
 DB: 6 Gaps: 0  
 US-10-669-693-2 (1-504) x AX876954 (1-2513)  
 QY 1 MetGluAlaThrGlyThrTyrAlaLeuLeuLeuAlaLeuAlaLeuLeuLeuLeuThr 20  
 Db 56 ATGAGGCGACCGGACCTGGCGGTGCTGTGGCGCTGGCGCTGCTCTCTGCTGTGAGC 115  
 QY 21 LeuAlaLeuSerGlyThrArgAlaArgGlyHisLeuProGlyProThrProLeuPro 40  
 Db 116 CTGGCGCTGTCCGGACCGAGGCCCGACCTGCGCCCGGCGCCACCGCGGTACCA 175  
 QY 41 LeuLeuGlyAsnLeuLeuLeuArgProGlyAlaLeuTyrSerGlyLeuMetArgLeu 60  
 Db 176 CTGTGGGAAACCTCTCTGACGTACGGCCCGGGCGCTGTATTACGGGCTCATGGGCTG 235

QY	61	SerLysLysTyrGlyProValPheThrIleTyrLeuGlyProThrArgProValValVal	80
Db	236	AGTAAGAGTACGACCGGTGTTCCACATCTACCTGGAGCCCTGGCGCCCTGTGGTGC	295
QY	81	LeuValGlyGlnGluAlaValArgGluAlaLeuGlyGlyGlnAlaGluGluPheSerGly	100
Db	296	CTGGTGGGAGGAGGCTGTGGGAGGCCCTGGGAGGTGAGGAGTTTCAGCGGC	355
QY	101	ArgGlyThrValAlaMetLeuGluGlyThrPheAspGlyHisGlyValPhePheSerAsn	120
Db	356	CGGGAAACCGTAGCGATGCTGGAGGGACTTTTATGAGCCATGGGGTTTCTTCTCCAC	415
QY	121	GlyGluArgTyrArgGlnLeuArgLysPheThrMetLeuAlaLeuArgAspLeuGlyMet	140
Db	416	GGGAGCGGTGGAGGACGTGAGGAAGTTTACCATGCTTGTCTGCGGGACCTGGGCATG	475
QY	141	GlyLysArgGluGlyGluGluLeuGlnAlaGluAlaAraGlyLeuValGluThrPhe	160
Db	476	GGGAAGCGGAGGAGGAGTGCATCCAGCGGAGGCCCGGTGTCTGGTGGAGACATTC	535
QY	161	GlnGlyThrGluGlyArgProPheAspProSerLeuLeuLeuAlaGlnAlaThrSerAsn	180
Db	536	CAGGGGACAGAGGACGCCATTCGATCCCTCCCTGCTGCTGGCCAGCCACCTCCAC	595
QY	181	ValValCysSerLeuLeuPheGlyLeuArgPheSerTyrGluAspLysGluPheGlnAla	200
Db	596	GTAGTCTGCTCCTCTCTTTGGCCTCGCTTCTCTATGAGGATAAGGAGTTCAGGGC	655
QY	201	ValValArgAlaAlaGlyThrLeuLeuGlyValSerSerGlnGlyGlyGlnThrTyr	220
Db	656	GTGGTCCGGGAGCTGGTGGTACCTCTGGAGTTCAGTCCAGGGGGGTGAGACCTAC	715
QY	221	GluMetPheSerTrpPheLeuArgProLeuProGlyProHisLysGlnLeuLeuHis	240
Db	716	GAGATGTTCTCTGGTTCCTCGGGCCCTCCAGGCCCCCAACAGCAGCTCTCCACAC	775
QY	241	ValSerThrLeuAlaAlaPheThrValArgGlnValGlnGlnHisGlnGlyAsnLeuAsp	260
Db	776	GTCCAGACCTTGGCTGCTTACAGTCCGGCAGGTGCAGCAGCACCCAGGGGAACCTGGAT	835
QY	261	AlaSerGlyProAlaArgLeuValAspAlaPheLeuLeuLysMetAlaGlnGluGlu	280
Db	836	GCTTCGGGCCCCAGCTGACCTTGTGATGCTTCTGCTGAAGATGGCACAGGAGAA	895
QY	281	GlnAsnProGlyThrGluPheThrAsnLysAsnMetLeuMetThrValIleTyrLeuLeu	300
Db	896	CAAAACCCAGGACAGAAATTCACCAACAGAAATGCTGATGACAGTCAITTTATTTGCTG	955
QY	301	PheAlaGlyThrMetThrValSerThrThrValGlyTyrThrLeuLeuLeuMetLys	320
Db	956	TTTGTCTGGGACGATGACGGTCAACACCGCTCGGCTATACCTCTCTCTCTGATGAAA	1015
QY	321	TyrProHisValGlnLysTrpValArgGluGluLeuAsnArgGluLeuGlyValGlyGln	340
Db	1016	TACCCTCATGTCAAAAGTGGTACGTGAGAGCTGAATCGGAGCTGGGGGTGGCCAG	1075
QY	341	AlaProSerLeuGlyAspArgThrArgLeuProTyrThrAspAlaValLeuHisGluAla	360
Db	1076	GCACCAAGCTTGGGACCGTACCCGCTCCCTTACACCGCGGCTTCTGTCATGAGCG	1135
QY	361	GlnArgLeuLeuAlaLeuValProMetGlyIleProArgThrLeuMetArgThrThrArg	380
Db	1136	CAGCGGTGTGGCGCTGGTGGCCATGGGAATACCCGCACTTCATGCGGACACCCCGC	1195
QY	381	PheArgGlyTyrThrLeuProGlnGlyThrGluValPheProLeuLeuGlySerIleLeu	400
Db	1196	TTCCGAGGGTACACCTTGCCCGGACGAGGTTCTTCCCTTCTTGGCTCCATCTCTG	1255
QY	401	HisAspProAsnIlePheLysHisProGluGluPheAsnProAspArgPheLeuAspAla	420
Db	1256	CATGACCCCAACATCTTCAAGCACCAGAGAGTTCAACCCAGACCGTTTCTCTGATGCA	1315
QY	421	AspGlyArgPheArgLysHisGluAlaPheLeuProPheSerLeuGlyLysArgValCys	440
Db	1316	GATGAGCGGTTCAGAGCATGAGCGTTCCTGCCCTTCTCTTAGGAGCGGTGCTCTGC	1375
QY	441	LeuGlyGluGlyLeuAlaLysAlaGluLeuPheLeuPhePheThrIleLeuGlnAla	460
Db	1376	CTTGAGAGGCGCTGGCAAGAGCGAGCTCTCTCTTTCACACCATCTTCAAGCC	1435
QY	461	PheSerLeuGluSerProCysProAspThrLeuSerLeuLysProThrValSerGly	480
Db	1436	TTCTCCCTGGAGAGCCCGTGCCTGCCCGGACACCTGAGCCTCAAGCCACCGTCAGTGC	1495
QY	481	LeuPheAsnIleProProAla	487
Db	1496	CTTTTCAACATTCCTCCCGACCC	1516
RESULT 11			
BD156406			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
<p>BD156406 2513 bp DNA linear PAT 17-JAN-2003</p> <p>Primer for synthesizing full-length cDNA and use thereof.</p> <p>BD156406</p> <p>BD156406.1 GI:27862164</p> <p>JP 2002191363-A/11249.</p> <p>Homo sapiens (human)</p> <p>Homo sapiens</p> <p>Eukaryota; Metazoa; Chordata; Vertebrata; Ruteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.</p> <p>1 (bases 1 to 2513)</p> <p>Otsu, T., Isogai, T., Nishikawa, T., Hayashi, K., Saito, K., Yamamoto, J., Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K. and Otsuki, T.</p> <p>Primer for synthesizing full-length cDNA and use thereof</p> <p>Patent: JP 2002191363-A 11249 09-JUL-2002;</p> <p>HELIX RESEARCH INSTITUTE</p> <p>OS Homo sapiens (human)</p> <p>PN JP 2002191363-A/11249</p> <p>PD 09-JUL-2002</p> <p>PF 28-JUL-2000 JP 2000280990</p> <p>PI TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU SAITO,</p> <p>PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,</p> <p>PI KEIICHI NAGAI, TETSUJI OTSUKI</p> <p>PC</p> <p>C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/10, C12P21/02, C12P21/68//C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC</p> <p>Primer for synthesizing full-length cDNA and use thereof FH key</p> <p>Location/Qualifiers (56) (1747).</p> <p>FT CDS</p> <p>Location/Qualifiers</p> <p>source</p> <p>1..2513</p> <p>/organism="Homo sapiens"</p> <p>/mol_type="genomic DNA"</p> <p>/db_xref="taxon:9606"</p>			
ORIGIN			
<p>Alignment Scores:</p> <p>Pred. No.: 6.2e-194 Length: 2513</p> <p>Score: 2527.00 Matches: 487</p> <p>Percent Similarity: 100.00% Conservative: 0</p> <p>Best Local Similarity: 100.00% Mismatches: 0</p> <p>Query Match: 96.63% Indels: 0</p> <p>DB: 6 Gaps: 0</p> <p>US-10-669-693-2 (1-504) x BD156406 (1-2513)</p>			
QY	1	MetGluAlaThrGlyThrTrpAlaLeuLeuAlaLeuAlaLeuLeuLeuLeuLeuThr	20
Db	56	ATGAGGCGGACCGGACCTGGCGCTGCTGTGGCGCTGGCGCTCTCTGCTGCTGACG	115
QY	21	LeuAlaLeuSerGlyThrArgAlaArgGlyHisLeuProGlyProThrProLeuPro	40
Db	116	CTGGCGCTGTCCGGGACCGAGGCCCGGAGCCACCTGCCCCCGGGCCACCGCTACCA	175



QY 41 LeuLeuGlyAsnLeuLeuGlnLeuArgProGlyAlaLeuTyrSerGlyLeuMetArgLeu 60  
DB 176 CTGCTGGAAACCTCTCTCAGCTACGGCCCGGGCGCTGTTATTCAGGCTCATGGCGTG 235  
QY 61 SerLysLysTyrGlyProValPheThrIleTyrLeuGlyProTrrArgProValVal 80  
DB 236 AGTAAGAAGTACGACCGGTGTTCACCATCTACCTGGGACCCCTGGCGGCTGTGGTGTC 295  
QY 81 LeuValGlyGlnGluAlaValArgGluAlaLeuGlyGlyGlnAlaGluGluPheSerGly 100  
DB 296 CTGCTGGGACGAGGCTGTGGGAGGCCCTGGAGGTCTAGGCTGAGAGTTTCAGCGGC 355  
QY 101 ArgGlyThrValAlaMetLeuGluGlyThrPheAspGlyHisGlyValPhePheSerAsn 120  
DB 356 CGGGAAACCGTAGCATGCTGGAAGGACTTTTTCATGCCATGGGTTCCTTCCTCAAC 415  
QY 121 GlyGluArgTrrArgGlnLeuArgLysPheThrMetLeuAlaLeuArgAspLeuGlyMet 140  
DB 416 GGGGAGCGGTGGAGGACCTGAGGAAGTTTACCATGCTTGTCTGCGGACCTGGGCATG 475  
QY 141 GlyLysArgGluGlyGluLeuLeuGlnAlaGluAlaArgCysLeuValGluThrPhe 160  
DB 476 GGGAGCGAGAGAGGAGGAGCTGATCCAGCGGAGGCCCGGTGTCTGGAGACATTC 535  
QY 161 GlnGlyThrGluGlyArgProPheAspProSerLeuLeuAlaGlnAlaThrSerAsn 180  
DB 536 CAGGGACAGAGGACGCCATTCGATCCCTCCCTGCTGCTGGCCAGGCCACCTCCAAC 595  
QY 181 ValValCysSerLeuLeuPheGlyLeuArgPheSerTyrGluAspLysGluPheGlnAla 200  
DB 596 GTAGTGTCTCCCTCTCTTGGGCTCCGCTCTCTATGAGGATAAGGAGTTCCAGGCC 655  
QY 201 ValValArgAlaAlaGlyGlyThrLeuLeuGlyValSerGlnGlyGlyGlnThrTyr 220  
DB 656 GTGCTCGGGCAGCTGTGTGTACCTCTGCTGGAGTCAGTCCCGAGGGGGTTCAGACCTAC 715  
QY 221 GluMetPheSerTrpPheLeuArgProLeuProGlyProHisLysGlnLeuLeuHisHis 240  
DB 716 GAGATGTCTCTCTGTTCTCGGCCCCCTGCCAGCCCCCACAGCAGCTCTCCACAC 775  
QY 241 ValSerThrLeuAlaAlaPheThrValArgGlnValGlnGlnHisGlnGlyAsnLeuAsp 260  
DB 776 GTCAGCACCTTGTGCTGCTTTCACAGTCCGCGAGTGCAGCAGCAGCAGGGAACCTGGAT 835  
QY 261 AlaSerGlyProAlaArgAspLeuValAspAlaPheLeuLeuLysMetAlaGlnGluGlu 280  
DB 836 GCTTCGGGCCCCGACGTGACCTTTCGATGCTCTCTGCTGAGATGGCAGCAGGAGGAA 895  
QY 281 GlnAsnProGlyThrGluPheThrAsnLysAsnMetLeuMetThrValIleTyrLeuLeu 300  
DB 896 CAAACCCAGGCACAGATTCACCAACAGACATGCTGATGACAGTCATTTATTTGCTG 955  
QY 301 PheAlaGlyThrMetThrValSerThrThrValGlyTyrThrLeuLeuLeuMetLys 320  
DB 956 TTTGCTGGGACGATGACGTCAGCACCCAGCTGCTGCTATACCTCTCTCTCTGATGATA 1015  
QY 321 TyrProHisValGlnLysTrpValArgGluGluLeuAsnArgGluLeuGlyAlaGlyGln 340  
DB 1016 TACCCTCATGTCCAAAGTGGGTACGTGAGGAGCTGAATCGGGAGCTGGGGCTGGCCAG 1075  
QY 341 AlaProSerLeuGlyAspArgThrArgLeuProTyrThrAspAlaValLeuHisGluAla 360  
DB 1076 GCACCAAGCCTAGGGACCGTACCCGCCCTCCCTTACACCGACCGGTCTGCGATGAGGCG 1135  
QY 361 GlnArgLeuLeuAlaLeuValProMetGlyIleProArgThrLeuMetArgThrThrArg 380  
DB 1136 CAGCGGCTGTCGCTGGTGGTCCCATGGGAATACCCCGCACCTCATGCGGACCCCGC 1195  
QY 381 PheArgGlyTyrThrLeuProGlnGlyThrGluValPheProLeuLeuGlySerIleLeu 400  
DB 1196 TTCCCGGGGTACACCTGCCCGAGGACGAGGAGTCTTCCCTCTCTGCTCCATCTCTG 1255  
QY 401 HisAspProAsnIlePheLysHisProGluGluPheAsnProAspArgPheLeuAspAla 420

DB 1256 CATGACCCCAACAICTTCAAGCACCACCAAGAGTTCAACCCAGACCGTTTCTCGATGCA 1315  
QY 421 AspGlyArgPheArgLysHisGluAlaPheLeuProPheSerLeuGlyLysArgValCys 440  
DB 1316 GATGACGTTTCAGGACCATGAGCGTCTCTGCGCTTCTCTTAGGAAGCGTGTCTGCG 1375  
QY 441 LeuGlyGluGlyLeuAlaLysAlaGluLeuPheLeuPhePheThrIleLeuGlnAla 460  
DB 1376 CTTGGAGAGGCGCTGGCAAAAGCGAGCTCTCTCTCTTCCACCCATCCTCAAGGCC 1435  
QY 461 PheSerLeuGluSerProCysProAspProAspThrLeuSerLeuLysProThrValSerGly 480  
DB 1436 TTTCTCCCTGGAGAGCCGTGCGCGGACACCTTGAGCTCAAGCCACCGTCAGTGGC 1495  
QY 481 LeuPheAsnIleProProAla 487  
DB 1496 CTTTCAACATTCCCCAGGCC 1516  
RESULT 12  
AK027605 2513 bp mRNA linear PRI 01-AUG-2002  
LOCUS Homo sapiens cDNA FLJ14699 fis, clone NT2RP2006571, moderately  
DEFINITION similar to CYTOCHROME P450 2G1 (BC 1.14.14.1).  
ACCESSION AK027605  
VERSION AK027605.1 GI:14042395  
KEYWORDS oligo capping; fis (full insert sequence).  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1  
AUTHORS Isogai, T., Ota, T., Hayashi, K., Sugiyama, T., Otsuki, T., Suzuki, Y.,  
Nishikawa, T., Nagai, K., Sugano, S., Ishibashi, T., Fujimori, K.,  
Tanai, H., Kimata, M., Watanabe, M., Hiraoka, S., Ishii, S., Kawai, Y.,  
Saito, K., Yamamoto, J., Wakamatsu, A., Nakamura, Y., Nagahari, K.,  
Masuho, Y., and Kanehori, K.  
TITLE NEDO human cDNA sequencing project  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 2513)  
AUTHORS Isogai, T. and Otsuki, T.  
TITLE Direct Submission  
JOURNAL Submitted (10-MAY-2001) Takao Isogai, Helix Research Institute,  
Genomics Laboratory, 1532-3 Yana, Kisarazu, Chiba 252-0812, Japan  
(E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)  
COMMENT NEDO human cDNA sequencing project supported by Ministry of  
Economy, Trade and Industry of Japan; cDNA full insert sequencing;  
Research Association for Biotechnology; cDNA library construction;  
5'- & 3'- end one pass sequencing and clone selection; Helix  
Research Institute (supported by Japan Key Technology Center etc.)  
and Department of Virology, Institute of Medical Science,  
University of Tokyo.  
FEATURES  
Location/Qualifiers  
1..2513  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="NT2RP2006571"  
/cell\_line="NT2"  
/cell\_type="teratocarcinoma"  
/clone\_lib="NT2RP2"  
/note="cloning vector: pME18SFL3-mRNA from NT2 neuronal  
precursor cells after 2-weeks retinoic acid (RA)  
induction."  
56..1750  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="BAB55227.1"  
/db\_xref="GI:14042396"  
/translation="MEATGTWALLALALLLTALSGTRARGHLPGCTPLILGN  
LLQLRPGALYGLMFLSKYGPVFTIYLPWPVPVVLVQAEAVRALGQAESEPSGR  
TVAMLEGTFDGHGVFSGNGERWRQRKFTMLALRDLGMKREGELIQAEARCLVET  
QGTGRPFDPDLLLAQAATSNVVCSLFLGLRFSYEDKRFQAVVRAAGTLLGVSSQGG





/note="the first two nucleotides (AT) of the initiation codon are missing from this clone"

```

/codon_start=1
/product="cytochrome P-450 2B-Bx"
/protein_id="AA31432.1"
/db_xref="GI:404777"
/tran_table="EPSSLILAFLAGLILLFRGPKVAGRLPGSPDLPVLGNLLQ
MDRGLRSLRFLRGKRGVDFVFLQSDTAIRALDQDAEAFVSGRGKIAV
VDPIFGQYVPIFANGERWAKRFLSLATMRDFGMGKRSVERIQEACRVLSELRK
GALLNLTALPHSTNISI CSIVFGKRGKQVFLRLDILDFQSFSLISFSSQVPEL
DSGFLHFPCTHROIYRNLEQNTFTQSVKRAHRTDPSNPDFIDVLLRMKKOKS
PSPGELHPCQMLILVLSFPAGTETSTSLFGLMLKYDHVTVQKEIQEIVGISH
RPALDADKMPYTDIAHEIIRGLDILFGVPHVTVKDFQPGYVPKNTEVPGVLS
SALHDPDYFTPTFNPHGLDQALAKRNGFMSLGKRGVLGEGIGATELFLPFT
TILQNFSAISGVPPEIDILTPRESGVGNVPYSYQIRFLAR"
2007. . 2012

```

polyA\_signal  
ORIGIN

Alignment Scores:		
Pred. No.:	3.1e-96	Length:
Score:	1313.00	Matches:
Percent Similarity:	71.05%	Conservative:
Best Local Similarity:	50.72%	Mismatches:
Query Match:	50.21%	Indels:
DB:	4	Gaps:
	4	3
		2026
		247
		99
		137
		4
		3

US-10-669-693-2 (1-504) x RABP450BX (1-2026)

Qy	7	TrpAlaLeuLeuLeuAlaLeuLeuLeu-----LeuLeuLeuThrLeuAlaLeuSer	24
Db	5	TTTCAGCCTGTCTCTCTCTGGCTTTCTCTGCAGGCCTCTCTGCTCTCTTGTTCAGGGGC	64
Qy	25	GlyThrArgAlaArgGlyHisLeuProProGlyProThrProLeuProLeuLeuGlyAsn	44
Db	65	CACCCCAAGGCCACCGCGCGCTCCCCCAGAGCCCTCCCTCTGCGCGCTCTGGGGAC	124
Qy	45	LeuLeuGlnLeuAArgProGlyAlaLeuTyrSerGlyLeuMetArgLeuSerLysIleTyr	64
Db	125	CTTCTGCAGATGGACAGGAAGGCCGTGCTCGGCTCCTTCTCTGCGGCTCCGAGAGAAATAC	184
Qy	65	GlyProValPheThrIleTyrLeuGlyProItrPargProValValLeuValGlyGln	84
Db	185	GGGACCGTGTTCACGGTGACTCGGATCC---AGACCCGTGCTCTGCTGTGGGACG	241
Qy	85	GluAlaValArgGluAlaLeuGlyGlyGlnAlaGluPheSerGlyArgGlyThrVal	104
Db	242	GATGCCATCCGGAGGCCCTCTGTGCACCAAGCTGAGGCCTTTCTGGCAGGGGAAGATC	301
Qy	105	AlaMetLeuGluGlyThrPheAspGlyHisGlyValPhePheSerAsnGlyGluArgTrp	124
Db	302	GCGTGTGGATCCGATCTTCAGGGATACGGAGTGATCTTTGCCAACGGGAGGCGCTGG	361
Qy	125	ArgGlnLeuArgGlyPheThrMetLeuAlaLeuArgAspLeuGlyMetGlyIysArgGlu	144
Db	362	CGGGCCCTTCGGAGATTCTCCCTGGCCACCATGTCGGGACTTCGGCATGGGAAGCGGAGC	421
Qy	145	GlyGluGluLeuIleGlnAlaGluAlaArgCysLeuValGluThrPheGlnGlyThrGlu	164
Db	422	GTGGAGGAGCGCATTCAGAGGAGGAGCCCGGTGTCTGTGGAGGAGCTCGGAAATCCCAAG	481
Qy	165	GlyArgProPheAspProSerLeuLeuAlaGlnAlaThrSerAsnValValCysSer	184
Db	482	GGAGCCCTCTGGACAAACACCTTGCTGTCTCCACTCAATCACCTCCACATCATCTGCTCC	541
Qy	185	LeuLeuPheGlyLeuArgPheSerTyrGluAspLysGluPheGlnAlaValValArgAla	204
Db	542	ATTGTCTTTGGAAAACGGCTTCACATACAGGACCCCGTGTCTCTCGCGCTGTGGACCTG	601
Qy	205	AlaGlyGlyThrLeuLeuGlyValSerSerGlnGlyGlyGlnThrThrGluMetPheSer	224
Db	602	TTCTTCCAGTCTCTCTCCCTCATCAGTCTCTCTCCAGCCAGGTGTTCAGGCTCTTCTCG	661
Qy	225	TrpPheLeuArgProLeuProGlyProHisLysGlnLeuLeuHisHisValSerThrLeu	244

662	GGCTTCCTAAAGCACTTTCTTGGCAGCGCAGCAGCATCTACAGAACTCTGCAGGAGAT	721
245	AlaAlaPheThrValArgGlnValGlnGlnHisGlnGlyAsnLeuAspAlaSerGlyPro	264
722	AACACTTTTCATCCACGAGCGTAGAAGACACCGCGCAACCTTGACCCCGACCAACCCC	781
265	AlaArgAspLeuValAspAlaPheLeuLeuLeuLysMetAlaGlnGluGlnAsnProGly	284
782	---AGGATTTCATCGACGTCTACCTGCTCGGATGGAAAAAGACACAGTCCGACCAAGC	838
285	ThrGluPheThrAsnLysAsnMetLeuMetThrValIleTyrIleuLeuPheAlaGlyThr	304
839	AGCGAGTTCACACGACGACCAACCTCATCTCTCACGGTCTCTCGCTCTTCTTCGCGGCAC	898
305	MetThrValSerThrThrValGlyTyrThrLeuLeuLeuMetLysTyrProHisVal	324
899	GAGACCACGACGACCCCTCGCTACGGCTTCTGCTCATGTCAAGTACCCACACGTC	958
325	GlnLysTrpValArgGluGluLeuAsnArgGluLeuGlyAlaGlyGlnAlaProSerLeu	344
959	ACAGAGAGATGTCAGAAAGAGATGTGACGAGTTCATCGGTCCACCGCTCCGCGCCCTC	1018
345	GlyAspArgThrArgLeuProTyrThrAspAlaValIleHisGluAlaGlnArgLeuLeu	364
1019	GATCACCAGCGCAAAATGCGCTACAGCGACGCGGTTCATCCACGAGATCCACGGGCTCGGG	1078
365	AlaLeuValProMetGlyIleProArgThrLeuMetArgThrThrArgPheArgGlyTyr	384
1079	GACCTCATCCCTTCGCGGGTGCCCCACACGGTCACAAAAGACACACAGTTCGAGGCTAT	1138
385	ThrLeuProGlnGlyThrGluValPheProLeuLeuGlySerIleuHisAspProAsn	404
1139	GTCAATCCCAAGAACACCGAAGTGTCCCGTCTGTAGCTCGGTCTTCCATGACCCGCGC	1198
405	IlePheLysHisProGluGluPheAsnProAspArgPheLeuAspAlaAspGlyArgPhe	424
1199	TACTTTAAACACCGAACACCTTCAACCCCGGCCACTTCTGATGTCACACGGGGCACTG	1258
425	ArgLysHisGluAlaPheLeuProPheSerLeuGlyLysArgValCysLeuGlyGluGly	444
1259	AAGAGGAATGAAGGCTTTATGCCCTCTCTCCCTGGGGAAGCGGTTGTCTGGGGCAAGGC	1318
445	LeuAlaLysAlaGluLeuPheLeuPheThrThrIleLeuGlnAlaPheSerLeuGlu	464
1319	ATCGCGCGGACCGAGCTGTCTCTTTTCACCACTCATCTGCAAGAACTTCTCCATGCC	1378
465	SerProCysProProAspThrLeuSerLysProThrValSerGlyLeuPheAsnIle	484
1379	AGCCCGTGCCTCCCGAGGACATCGACCTCACTCCCGGAGAGTGGCGTGGGCAAGGTG	1438
485	ProProAlaPheGlnLeuGln	491
1439	CCCCCGAGTACCATCCG	1459

RESULT 15

LOCUS	S64259	2028 bp	mRNA	linear	MAM 23-SEP-1993
DEFINITION	cytochrome P450 2B-Bx-phenobarbital-inducible [rabbits, kidney, mRNA, 2028 nt].				

ACCESSION

**VERSION**

## KEYWORDS

**SOURCE**

## ORGANISM

## REFERENCE

## AUTHORS

**TITLE**

TOYENSTAY

JOURNAL



**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 18, 2004, 00:56:16 ; Search time 123 Seconds  
(without alignments)  
2273.945 Million cell updates/sec

Title: US-10-669-693-2

Perfect score: 2615

Sequence: 1 MEATGTWALLALALLLT.....PPAFQVQRPTDLHSTTQTR 504

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO.spool/US10669693/runat\_15092004\_101311\_5876/app.query.fasta\_1.647  
-DB=Issued Patents NA -QWMT=fastcap -SUFFIX=rn1 -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=Dits -START=1 -END=1 -MATRIX=Blosum62 -TRANS=human40.cdi  
-LIST=45 -LOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10669693 @CGN 1.1.69 @runat\_15092004\_101311\_5876 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSBLOK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA:  
1: /cgn2\_6/ptodata/2/ina/5A.COMB.seq:\*  
2: /cgn2\_6/ptodata/2/ina/5B.COMB.seq:\*  
3: /cgn2\_6/ptodata/2/ina/6A.COMB.seq:\*  
4: /cgn2\_6/ptodata/2/ina/6B.COMB.seq:\*  
5: /cgn2\_6/ptodata/2/ina/PTCUS.COMB.seq:\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1238	47.3	1476	4	US-08-277-031B-10
2	1232	47.1	2907	4	US-09-023-655-1053
3	1213.5	46.4	1485	4	US-08-277-031B-8
4	1213.5	46.4	1485	4	US-08-277-031B-9
5	1185	45.3	1737	2	US-08-750-703-2
6	1175.5	45.0	1740	2	US-08-750-703-1
7	1108	42.4	1825	4	US-09-023-655-1061
8	1087.5	41.6	1854	1	US-08-201-118-4
9	1087.5	41.6	1854	2	US-08-238-821B-4
10	1087.5	41.6	1854	4	US-09-023-655-1056
11	1087.5	41.6	1854	5	PCT-US95-05744-4
12	1086.5	41.5	1473	4	US-08-277-031B-2

13	1085.5	41.5	1473	4	US-08-277-031B-15
14	1085.5	41.5	1845	4	US-09-023-655-1057
15	1085.5	41.5	1852	1	US-08-201-118-10
16	1085.5	41.5	1852	2	US-08-238-821B-10
17	1085.5	41.5	1852	5	PCT-US95-05744-10
18	1084.5	41.5	1746	1	US-08-201-118-2
19	1084.5	41.5	1746	2	US-08-238-821B-2
20	1084.5	41.5	1746	4	US-09-023-655-1059
21	1084.5	41.5	1746	5	PCT-US95-05744-2
22	1078.5	41.2	1591	4	US-08-194-981E-3
23	1073.5	41.1	1473	4	US-08-277-031B-14
24	1073.5	41.1	1995	4	US-09-023-655-1055
25	1073.5	41.1	2009	1	US-08-201-118-6
26	1073.5	41.1	2009	2	US-08-238-821B-6
27	1073.5	41.1	2009	5	PCT-US95-05744-6
28	1067.5	40.8	2258	1	US-08-201-118-12
29	1067.5	40.8	2258	2	US-08-238-821B-12
30	1067.5	40.8	2258	4	US-09-023-655-995
31	1067.5	40.8	2258	5	PCT-US95-05744-12
32	1059	40.5	1419	2	US-08-194-981E-4
33	1054.5	40.3	1473	4	US-08-277-031B-11
34	1051.5	40.2	1473	4	US-08-277-031B-12
35	1050.5	40.2	1829	1	US-08-201-118-8
36	1050.5	40.2	1829	2	US-08-238-821B-8
37	1050.5	40.2	1829	5	PCT-US95-05744-8
38	1044.5	39.9	1473	4	US-08-277-031B-13
39	1033.5	39.5	1892	1	US-08-201-118-14
40	1033.5	39.5	1892	2	US-08-238-821B-14
41	1033.5	39.5	1892	5	PCT-US95-05744-14
42	982.5	37.6	1482	2	US-08-194-981E-1
43	982.5	37.6	1482	4	US-08-277-031B-3
44	960.5	36.7	1419	2	US-08-194-981E-2
45	871.5	33.3	1857	4	US-09-023-655-902

#### ALIGNMENTS

##### RESULT 1

US-08-277-031B-10  
; Sequence 10, Application US/08277031B  
; Patent No. 6620593  
; GENERAL INFORMATION:  
; APPLICANT: Hayashi, Koji  
; APPLICANT: Sakaki, Toshiyuki  
; APPLICANT: Yabusaki, Yoshiyasu  
; APPLICANT: Komai, Koichiro  
; APPLICANT: Kaneko, Hideo  
; APPLICANT: Nakatsuka, Iwao  
; TITLE OF INVENTION: METHOD FOR SAFETY EVALUATION OF  
; TITLE OF INVENTION: CHEMICAL COMPOUND USING RECOMBINANT YEAST EXPRESSING  
; TITLE OF INVENTION: HUMAN CYTOCHROME P450  
; NUMBER OF SEQUENCES: 42  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Birch, Stewart, Kolasch & Birch  
; STREET: P.O. Box 747  
; CITY: Falls Church  
; STATE: Virginia  
; COUNTRY: USA  
; ZIP: 22040-0747  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.5inch, 1.44MB  
; COMPUTER: IBM PC  
; OPERATING SYSTEM: Dos 5.0  
; SOFTWARE: WordPerfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/277,031B  
; FILING DATE: 19-JULY-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP-201120/1993  
; APPLICATION NUMBER: JP-180246/1993  
; APPLICATION NUMBER: JP-208279/1993  
; FILING DATE: 20-07-1993

FILING DATE: 21-07-1993  
 FILING DATE: 30-07-1993  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Raymond C. Stewart  
 REGISTRATION NUMBER: 21,066  
 REFERENCE/DOCKET NUMBER: 20-3530P  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 703-205-8000  
 TELEFAX: 703-205-8050  
 TELEFAX:

INFORMATION FOR SEO ID NO: 10:

; INFORMATION FOR SEQ 1:
: SEQUENCE CHARACTERISTICS:

```

; SEQUENCE CIPHERTEXT:
; LENGTH: 1476

```

LENGTH. 1470  
TYPE: nucleic acid

TYPE: nucleic acid  
STRANDEDNESS: double

STRANDEDNESS: 000  
TOPOLOGY: linear

TOPLOGY:  
; 118-08-277-031B-10

Alignment Scores:

Alignment Scores:	
Pred. No.:	3.39e-125

FILED: NO.:	238
SCORE:	1238.00
MATCHES:	238

percent similarity: 68.51% Conservative: 97

Percent Similarity:	98.52%	98.52%
Best Local Similarity:	48.67%	150
Mismatches:		

BEST LOCAL SIMILARITY:	49.07%	4
Overall Match:	47.34%	4
Indels:		

Query Match:	4 / 346	3
nr.	4	3
Gaps:		
Inserts:		

DB:                    4                    caps.

US-10-669-693-2 (1-504) x US-08-277-031B-10 (1-1476)

Qy	8	AlaLeuLeuLeuAlaLeuAlaLeu-----LeuLeuLeuThrLeuAlaLeuSerGly	25
Db	10	AGCGTCCTCTCTTCCTGCACTCTCTCAAGAGACTCTTGCTACTCTCGTTCAGCGCCAC	69
Qy	26	ThrArgAlaArgGlyHisLeuProProGlyProThrProLeuProLeuLeuGlyAsnLeu	45
Db	70	CCTAACACCCATGACCGCTCTCCACCAGGCCCCCGCTCTGCCCTTTTGGAAACCTT	129
Qy	46	LeuGlnLeuArgProGlyAlaLeuTyrSerGlyLeuMetArgLeuSerIysIysTyrGly	65
Db	130	CTGCAGATGGATAGAAGAGGCGCTACTCAAAATCCTTCTGAGGTTCGAGAGAAATATGGG	189
Qy	66	ProValPheThrIleTyrLeuGlyProIrrArgProValValValLeuValGlnGlu	85
Db	190	GACGTCCTTCAGCGTATCACTGGGACG---AGGCCCGTGTCTATGCTGTGTGAGTAGAG	246
Qy	86	AlaValArgGluAlaLeuGlyGlyGlnAlaGluGluPheSerGlyArgGlyThrValAla	105
Db	247	GCCATACGGGAGGCCCTTGTGGCAAGGCTGAGGCCCTTCTTGCGCGGGGAAAAATCGCC	306
Qy	106	MetLeuGluGlyThrPheAspGlyHisGlyValPhePheSerAsnGlyGluArgTrpAsp	125
Db	307	ATGGTCGACCCCATCTTCGGGGATATGTGTGATCTTTGCCAATGGAAACCGCTGGAAG	366
Qy	126	GlnLeuArgLysPheThrMetLeuAlaLeuArgAspLeuGlyMetGlyLysArgGluGly	145
Db	367	GTGCTTCGGCGATTCTCTGTGACCACCTATATGAGGACTTCGGGATGGAAAGCGAGTGTG	426
Qy	146	GluGluLeuIleGlnAlaGluAlaArgCysLeuValGluThrPheGlnGlyThrGluGly	165
Db	427	GAGGACGGGATTCAGAGAGGCGCTCAGTGTCTGATGAGAGGCTTCGGAAATCCAGGGG	486
Qy	166	ArgProPheAspProSerLeuLeuAlaGlnAlaThrSerAsnValValCysSerLeu	185
Db	487	GCCCTCATGACCCCACTTCTCTTCAGTCCATTACGCCCAACATCATCTGCTGCATC	546
Qy	186	LeuPheGlyLeuArgPheSerTyrGluAspLysGluPheGlnAlaValValArgAlaAla	205
Db	547	GTCTTTGGAAAAAGATTCCACTACCAAGATCAAGAGTTCCTGAGATGCTGAACTGTTC	606
Qy	206	GlyGlyThrLeuLeuGlyValSerSerGlnGlyGlnThrThrTyrGluMetPheSerTrp	225
Db	607	TACCAGACTTTTTCACTCATCAGCTCTGTATTCCGCCGCGTGTGTGAGCTTCTCTCGGC	666
Qy	226	PheLeuArgProLeuProGlyProHisLysGlnLeuLeuHisHisValSerThrLeuAla	245

D	b		667	TTCCTGAATAACTTTCCCTGGGCGCACACAGGCAAGTTTTCAA AAAAACCCTGCAGAAATCAAT	726
Q	y		246	AlaPheThrValArgGlnValGlnGlnHisGlnGlyAsnLeuAspAlaSerGlyProAla	265
D	b		727	GUTTATCATTTGGGCCACAGTGGTGAGAAGACACCGTGAAACCTTCGGACCCCAGCGCCCC--	783
Q	y		266	ArgASPLeuValAspAlaPheLeuLeuLysMetAlaGlnGluGluGlnAsnProGlyThr	285
D	b		784	ARGGACCTCATTGCACACCTTA CTGCTCCACATGGAAAAAGAAAATCCAACGCACACAGT	843
Q	y		286	GlupHeThrAsnLysAsnMetLeuMetThrValIleTy rLeuLeuPhealaglyThrMet	305
D	b		844	GAATTCAGGCACCCAGAACCTCAACTCAACAGCTCTCGCTCTTCTTCTGGCACTGAG	903
Q	y		306	ThrValSerThrThrValGlyTy rThrLeuLeuLeuLeuMetLysTy rProHisValGln	325
D	b		904	ACCACAGCACCACCACTCTCCGCTACGGCTTCCTGCTCATGCTCAATACCTCATGTGCA	963
Q	y		326	LvsTrpValArgGluGluLeuAsnArgGluLeuLeuGlyAlaGlyGlnAlaProSerLeuGly	345
D	b		964	GAGAGAGTCTACAGGGAGATTGAACAAGGTGATTGGGCCACATCGCCCTCCAGAGCTTCAT	1023
Q	y		346	AspArgThrArgLeuProTy rThrAspAlaValLeuHisGluAlaGlnArgLeuLeuAla	365
D	b		1024	GACCGAGCCAAAATGCATACACAGAGCGAGT CATCTATGAGATT CAGAGATTTCCGCAC	1083
Q	y		366	LeuValProMetGlyIleProArgThrLeuMetArgThrThrArgPheArgGlyTy rThr	385
D	b		1084	CITTCCTCCCATGGGTGTGGCCCCACAT TGTCACCCACACACACAGCTTCCAGGGTACATC	1143
Q	y		386	LeuproGlnGlyThrGluValPheProLeuLeuGlySeri leLeuHisaspProAsnile	405
D	b		1144	ATCCCCAAGGACACAGAAGTATTCTCATCTCCTGAGCACTCTCTCCATGACCCACACATC	1203
Q	y		406	PheLyshisProGluGluPu heAsnProAspArgPheLeuAspAlaaspGlyArgPheArg	425
D	b		1204	TTTGAAAAACAGAGCGCTTCAATCTCTGACCACTTCTTGATGCCAATGGGCGACTGAAA	1263
Q	y		426	LysHisiGluAlaPheLeuProPheSerLeuGlyl ysaArgVal Cysl eugGlyGluGlyLeu	445
D	b		1264	AAGACTGAAGCTTTTATCCCTCTCTCCTTAGGGAAGCGGATTTGTCTTGTTGAAGGCATC	1323
Q	y		446	AlalyalaGluLeuPheLeuPhePheThrThrIleLeuGlnAlaPheSerLeuGluSer	465
D	b		1324	GCCCGTGGGAATTTGTTCTCTCTTCACCACCAATCTCCAGAACTTCTCCATGGCCAGC	1383
Q	y		466	ProcysProAspThrLeuSerLeuLysProThrValSerGlyLeuPheAsnilePro	485
D	b		1384	CCCGTGGCCCCAGAAGATCGATCTGACACCCCGAGGAGTGTGGTGTGGGCAAAATACCC	1443
Q	y		486	ProAlaPheGlnLeuGlnValArgpro	494
D	b		1444	CTACATACCCAGATCCGCTTCTCTGCC	1470

## RESULT 2

RESULTS

US-09-023-655-1033  
: Sequence 1053. Application US/09023655

: Patent No. 6607879  
: Sequence ID: 1033, App

; PALENC NO: 6607873  
: GENERAL INFORMATION:

APPLICANT: Cocks.

APPLICANT: SUSAN G. STUART  
APPLICANT: SUSAN G. STUART

APPLICANT: Susan G. Seegar  
APPLICANT: Jeffrey J. Seilhamer

APPLICANT:	INVENTOR:	TITLE OF INVENTION:	COMPOSITION:
APPLICANT: JELLEY S. SCRAMMER	INVENTOR: JELLEY S. SCRAMMER	TITLE OF INVENTION: COMPOSITION	COMPOSITION:

CONTRACT NO.	TITLE OF INVENTION:	EXPRESSION
1	TITLE OF INVENTION:	EXPRESSION

FILE OF INVENTION: EXTENSION  
NUMBER OF SEQUENCES: 1508

; ; NUMBER OF SEQUENCES: 1508  
; ; CORRESPONDENCE ADDRESS:; CORRESPONDENCE ADDRESS:  
ADDRESS: INCYTE PHARMACEUTICALS, INC.  
10000 BELLEVUE AVENUE  
SUITE 100  
BELLEVUE, WA 98004  
USA  
TEL: 206 462 1000  
FAX: 206 462 1001  
E-MAIL: [info@incyte.com](mailto:info@incyte.com)  
WWW: [www.incyte.com](http://www.incyte.com)  
;

ADDRESS: INCYTE PHARMACEUTICALS  
CITY: 3174 PORTER DRIVE

STREET: 31/4 PORTER DRIVE  
CITY: DAYTON, OHIO

;  
CITY: PALO ALTO  
STATE: CALIFORNIA

STATE: CALIFORNIA  
COUNTRY: USA

; COUNTRY: USA  
; ZIP: 04304

; ZIP: 94304

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/023,655  
FILING DATE: HEREWITH  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Zeller, Karen J.  
REGISTRATION NUMBER: 37,071  
REFERENCE/DOCKET NUMBER: PA-0001 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (650) 855-0555  
TELEFAX: (650) 845-4166  
FORMATION FOR SEQ ID NO: 1053:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2907 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: GENBANK  
CLONE: g181293  
023-655-1053

```

Alignment Scores:
Pred. No.:      4.52e-124      Length:      2907
Score:          1232.00        Matches:     235
Percent Similarity: 68.71%      Conservative: 101
Best Local Similarity: 48.06%    Mismatches:  149
Query Match:    47.11%        Indels:      1
DB:             4             Gaps:        3
US-10-669-693-2 (1-504) x US-09-023-655-1053 (1-2907)

```

```

1 APPLICANT: Hayaashi, Koji
2 APPLICANT: Sakaki, Toshiyuki
3 APPLICANT: Yabasaki, Yoshiyasu
4 APPLICANT: Komai, Koichiro
5 APPLICANT: Kaneko, Hideo
6 APPLICANT: Nakatsuka, Iwao
7 TITLE OF INVENTION: METHOD FOR SAFETY EVALUATION OF
8 TITLE OF INVENTION: CHEMICAL COMPOUND USING RECOMBINANT YEAST EXPRESSING
9 TITLE OF INVENTION: HUMAN CYTOCHROME P450
10 NUMBER OF SEQUENCES: 42
11 CORRESPONDENCE ADDRESS:
12 ADDRESSEE: Birch, Stewart, Kolasch & Birch
13 STREET: P.O. Box 747
14 CITY: Falls Church
15 STATE: Virginia
16 COUNTRY: USA
17 ZIP: 22040-0747
18 COMPUTER READABLE FORM:
19 MEDIUM TYPE: Diskette, 3.5inch, 1.44MB
20 COMPUTER: IBM PC
21 OPERATING SYSTEM: Dos 5.0
22 SOFTWARE: WordPerfect 5.1
23 CURRENT APPLICATION DATA:
24 APPLICATION NUMBER: US/08/277,031B
25 FILING DATE: 19-July-1994
26 CLASSIFICATION: 435
27 PRIOR APPLICATION DATA:
28 APPLICATION NUMBER: JP-201120/1993
29 APPLICATION NUMBER: JP-180246/1993
30 APPLICATION NUMBER: JP-208279/1993
31 FILING DATE: 20-07-1993
32 FILING DATE: 21-07-1993
33 FILING DATE: 30-07-1993
34 ATTORNEY/AGENT INFORMATION:
35 NAME: Raymond C. Stewart
36 REGISTRATION NUMBER: 21,066
37 REFERENCE/DOCKET NUMBER: 20-3530P
38 TELECOMMUNICATION INFORMATION:
39 TELEPHONE: 703-205-8000
40 TELEFAX: 703-205-8050
41 TELEX:
42 INFORMATION FOR SEQ ID NO: 8:
43
44 SEQUENCE CHARACTERISTICS:
45 LENGTH: 1485
46 TYPE: nucleic acid
47 STRANDEDNESS: double
48 TOPOLOGY: linear
49
50 US-08-277-031B-8
51
52 Alignment Scores:
53 Pred. No.: 1,62e-122 Length: 1485
54 Score: 1213.50 Matches: 235
55 Percent Similarity: 68.38% Conservative: 98
56 Best Local Similarity: 48.25% Mismatches: 151
57 Query Match: 46.41% Indels: 3
58 DB: 4 Gaps: 3
59
60 US-10-669-693-2 (1-504) x US-08-277-031B-8 (1-1485)
61
62 QY 8 AlaLeuLeuAlaLeuAlaLeuLeuLeuLeuLeuThrLeuAlaLeuSerGlyThrArg 27
63 Db GCCTTGTGGTCGCCTGACTGATGGTCTTGATGTCGTGTT--TGGCAGCAGAGGAAG 84
64
65 QY 28 AlaArgGlyHisLeuProGlyProThrProLeuProLeuGlyAsnLeuLeuGln 47
66 Db AGCAAGGGGAAGCTGCCCTCCGGGACCCACCCTTCCTTTCATTGGAAACTACCTCGAG 14
67
68 QY 48 LeuArgProGlyAlaLeuTyrSerGlyLeuMetArgLeuSerIysLysTyrGlyProVal 67
69 Db CTGAACACAGACAGAGATGTAACAATCCCTCATGAAGATCAGTCAGCGCTATGCCCCGTG 20
70
71 QY 68 PheThrIleTyrLeuGlyProTfrrArgProValValLeuValGlyGlnGluAlaVal 87
72 Db

```



FEATURE:  
NAME/KEY: CYP2A13  
LOCATION:  
OTHER INFORMATION:  
US-08-750-703-2  
Alignment Scores:  
Pred. No.: 2,68e-119 Length: 1737  
Score: 1185.00 Matches: 241  
Percent Similarity: 66.06% Conservative: 88  
Best Local Similarity: 48.33% Mismatches: 160  
Query Match: 45.32% Indels: 10  
DB: 2 Gaps: 4  
US-10-669-693-2 (1-504) x US-08-750-703-2 (1-1737)  
QY 1 MetGluAlaThrGlyThrTrpAlaLeuLeuAlaLeuLeuLeuLeuLeuThr 20  
Db 1 ATGCTGGCCCTCAGG-----CTGCTTCTGGTGACCTTGTGCTGCTGCTGCTGCTG 51  
QY 21 LeuAlaLeuSerGlyThr-----ArgAlaAArgGlyHisLeuProGlyPro 36  
Db 52 ATGCTTCTGATGTCTGCTGCGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 111  
QY 37 ThrProLeuProLeuGlyAsnLeuLeuGlnLeuLeuLeuLeuLeuLeuLeuLeu 56  
Db 112 ACCCATTCCTTTCATTCGAACTACCTCCAGCTGAACACAGAGAGAGAGAGAGAG 171  
QY 57 LeuMetArgLeuSerLysLysTyrGlyProValPheThrIleTyrLeuGlyProTyrArg 76  
Db 172 CTGATGAAGATCAGTGGCGCTATGGCCCTGTGTTTCCATTCCTTCTGGGGCCC---CGG 228  
QY 77 ProValValLeuValGlyGlnGluAlaValAArgGluAlaLeuGlyGlyGlnAlaGlu 96  
Db 229 CGGCTCGTGTGTCTGTCGGACATGATGCGTCAAGAGAGCTCTGCTGGTGGACAGGCTGAG 288  
QY 97 GluPheSerGlyArgGlyThrValAlaMetLeuGluGlyThrPheAspGlyHisGlyVal 116  
Db 289 GAGTTCAGCGGGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAG 348  
QY 117 PhePheSerAsnGlyGluArgTyrArgGlnLeuArgLysPheThrMetLeuAlaLeuArg 136  
Db 349 CGCTTCAGCAACCGGGAGCGCGCAAGAGCTCGGGCGCTTCCATCGCCACCCCTAAGG 408  
QY 137 AspLeuMetGlyLysArgGlyGluGluLeuLeuGlnAlaGluAlaArgCysLeu 156  
Db 409 GGTCTTGGCGTGGCAAGCGCGCATCGAGGAGCGCATCGAGGAGGAGGAGGAGGAGGAG 468  
QY 157 ValGluThrPheGlnGlyThrGluGlyArgProPheAspProSerLeuLeuLeuAlaGln 176  
Db 469 ATCGAGCGCCCTCGGGGCGAGCGCGCATCGAGGAGGAGGAGGAGGAGGAGGAGGAG 528  
QY 177 AlaThrSerAsnValValCysSerLeuLeuPheGlyLeuArgPheSerTyrGluAspLys 196  
Db 529 ACAGTCTCCCAATGTCATCAGCTCCATTTGTTGGGAGCGCTTGGAGTATGAGGACAAA 588  
QY 197 GluPheGlnAlaValValArgAlaAlaGlyGlyThrLeuLeuGlyValSerSerGlnGly 216  
Db 589 GAGTTCCTGTCTGCTGTCGGCATGATGCTGGGAAGTTCAGTTCACGGGAACCTCCACG 648  
QY 217 GlyGlnThrTyrGluMetPheSerTrpPheLeuArgProLeuProGlyProHisLysGln 236  
Db 649 GGGCAGCTCTATGAGATGTTCTCTTCGGTGTATGAACACACCTGCGCAGGAGACA-CAGCAACAG 707  
QY 237 LeuLeuHisHisValSerThrLeuAlaAlaPheThrValArgGlnValGlnGlnHisGln 256  
Db 708 GCCTTTAAGGAGCTGCAAGGCTGGAGGACTTCATCGCCCAAGAGGTGGAGCACAACCCAG 767  
QY 257 GlyAsnLeuAspAlaSerGlyProAlaAArgAspLeuValAspAlaPheLeuLeuLysMet 276  
Db 768 CGCAGCTGGATCCCAATTCGCCA---CGGGAGCTTCATCGACTCCTCTTCTATCGCATG 824  
QY 277 AlaGlnGluGluGlnAsnProGlyThrGluPheThrAsnLysAsnMetLeuMetThrVal 296

1039 GCCAAGATGCCCTACATGGAGCAGTGATGATCACGAGATCCAAAGATTGGAGACGTGATC 1098  
QY 368 ProMetGlyIleProArgThrLeuMetArgThrThrArgPheArgGlyTyrThrLeuPro 387  
Db 1099 CCATGATGTTGGCCCGGAGAGTAAAGAGACACCAAGTTTCGGGATTTCTTCTCCCT 1158  
QY 388 GlnGlyThrGluValPheProLeuLeuGlySerIleLeuHisAspProAsnIlePheLys 407  
Db 1159 AAGGCCACCGAAGTACCTATGCTGGCTCTGTGCTGAGAGACCCAGATTCTTCTCC 1218  
QY 408 HisProGluGluPheAsnProAspArgPheLeuAspAlaAspGlyArgPheArgLysHis 427  
Db 1219 AACCCCGAGGACTTCAATCCCGAGCATTCTCGAATGAGAGAGGGGAGTTTAAAGAGAGT 1278  
QY 428 GluAlaPheLeuProPheSerLeuGlyLysArgValCysLeuGlyGluGlyLeuAlaLys 447  
Db 1279 GATGCTTTTGTGCCCTTTTCCATCGGAAGCGGAAGTGTTCGGAGAAGGCTGGCCAGA 1338  
QY 448 AlaGluLeuPheLeuPheThrThrIleLeuGlnAlaPheSerLeuGluSerProCys 467  
Db 1339 ATGGAGCTCTTCTCTTTCACCACTGTCATGCAAGACTTCGGCTCAAGTCTCTCCAG 1398  
QY 468 ProProAspThrLeuSerLeuLysProThrValSerGlyLeuPheAsnIleProProAla 487  
Db 1399 TCACCTAAGACATTGACGTGTCTCCCAACACGTTGGGCTTTGCCACGATCCCAAGAAC 1458  
QY 488 PheGlnLeuGlnValArgPro 494  
Db 1459 TACACCATGAGCTTCTGCCC 1479  
RESULT 5  
US-08-750-703-2  
Sequence 2, Application US/08750703  
Patent No. 5891633  
GENERAL INFORMATION:  
APPLICANT: Gonzalez, Frank J.; Idle, Jeffrey R.  
TITLE OF INVENTION: DEFECTS IN DRUG  
TITLE OF INVENTION: METABOLISM  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Morgan & Finnegan  
STREET: 345 Park Ave.  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10154-0053  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WordPerfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/750,703  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/07605  
FILING DATE: 16-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Dorothy R. Auth  
REGISTRATION NUMBER: 36,434  
REFERENCE/DOCKET NUMBER: 2026-4196PCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 758-4800  
TELEFAX: (212) 751-6849  
INFORMATION FOR SEQ ID NO. 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1737 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA





```

Db      682 AAACACCTCCGAGGACACACAGAGCCCTTTCAGTTGCTGCAAGGGCTGGAGGACTTC 741
Qy      248 ThrValArgGlnValGlnGlnHISGlnGlyAenLeuAspAlaSerGlyProAlaAlaArgAsp 267
Db      742 ATAGCAAGAGGTGGAGCACACACAGCGCAGCTGGATCCCAATTCGCCCA---CGGGAC 798
Qy      268 LeuValAspAlaPheLeuLeuLeuMetAlaGlnGlnGlnGlnAsnProGlyThrGluPhe 287
Db      799 TTCATTGACTCTCTTCATCGCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 858
Qy      288 ThrAsnLysAsnMetLeuMetThrValIleThrLeuLeuPheAlaGlyThrMetThrVal 307
Db      859 TACTTGAAGAACTGATGATGAGCAGCTTGAACCTTTCATTCAGGACACCGAGAGCGTC 918
Qy      308 SerThrThrValGlyThrLeuLeuLeuMetLysThrProHisValGlnLysTrp 327
Db      919 AGCACACCCCTGACTATGCTTCTTCTGCTCATGAAGCACCAGAGGTGGAGGCCAAG 978
Qy      328 ValArgGlnGlnLeuAsnArgGlnLeuGlyAlaGlyGlnAlaProSerLeuGlyAspArg 347
Db      979 GTCCATGAGGAGATTCACAGATGATCGCAAGACCGCAGCCCAAGTTTGAGGACCGG 1038
Qy      348 ThrArgLeuProThrThrAspAlaValLeuHisGlnAlaGlnArgLeuLeuAlaVal 367
Db      1039 GCCAAGATCCCTACATGAGGAGGAGTATCCACGAGATCCAAAGATTTCGAGACGTGATC 1098
Qy      368 ProMetGlyIleProArgThrLeuMetArgThrThrArgPheArgGlyThrThrLeuPro 387
Db      1099 CCATGAGTTGGCCCGCAGAGTCAAAAAGAGGACCAAGTTTCGGGATTTCTTCTCCCT 1158
Qy      388 GlnGlyThrGluValPheProLeuLeuGlySerIleLeuHisAspProAsnIlePheLys 407
Db      1159 AAGGGCATGAAGATGTTCCCTATGTTGGGCTCCGTCTGAGACCTCAGGTCTTCTCC 1218
Qy      408 HisProGlnGluPheAsnProAspArgPheLeuAspAlaAspGlyArgPheArgLysHis 427
Db      1219 AACCCCGGACATTCATCCCGACCTTCTCTGGTGAGAGGGGCGAGTTTAAGAAGCGT 1278
Qy      428 GluAlaPheLeuProPheSerLeuGlyLysArgValCysLeuGlyGlnGlyLeuAlaLys 447
Db      1279 GATGCTTTTGTGCGCTTCTCCATCAGAAAGCGGAAGCTGTTTCGGAAGAGCCCTGGCCAGA 1338
Qy      448 AlaGlnLeuPheLeuPheThrThrIleLeuGlnAlaPheSerLeuGluSerProCys 467
Db      1339 ATGGAGCTCTTCTCTCTTCCACACCGTATCGAGAACTTCGCGCTCAAGTCTCCAG 1398
Qy      468 ProProAspThrLeuSerLeuLysProThrValSerGlyLeuPheAsnIleProProAla 487
Db      1399 TCACCTAAGGACATGAGGTGTCCTCCCAACACGCTGGGCTTTCGACGATCCACAGAAC 1458
Qy      488 PheGlnLeuGlnValArgPro 494
Db      1459 TACACCATGAGCTTCTCTGCC 1479

```

RESULT 7

```

US-09-023-655-1061
; Sequence 1061, Application US/09023655
; Patent No. 6607879
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:

```

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 5.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; FILING DATE: US/09/023,655
; APPLICATION NUMBER:
; PRIORITY DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0001 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1061:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1825 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: g181357
; US-09-023-655-1061

```

```

Alignment Scores: 7.37e-111 Length: 1825
Pred. No.: 1108.00 Matches: 230
Score: 64.73% Conservative: 93
Percent Similarity: 46.09% Mismatches: 162
Best Local Similarity: 42.37% Indels: 14
Query Match: 4 Gaps: 6
DB:

```

US-10-669-693-2 (1-504) x US-09-023-655-1061 (1-1825)

```

Qy      1 MetGluAlaThrGlyThrTrpAlaLeuLeuAlaLeuAlaLeuLeuLeuLeuLeuLeu 19
Db      56 ATGAGCAGCATATAGCAGCAGCATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 115
Qy      20 ThrLeuAlaLeuSerGlyThrArgAlaArgGlyHisLeuProGlyProGlyProGlyPro 39
Db      116 ACCCTAAGC-----TCAAGAGATAAGGAAAGCTGCTCCGGGAGCCAGACCCCTC 166
Qy      40 ProLeuLeuGlyAsnLeuLeuGlnLeuArgProGlyAlaLeuTyrSerGlyLeuMetArg 59
Db      167 TCAATCTGGGAAACCTGCTGCTTGTCTCCAAAGACATGCTGACTTCTCTCACTAAG 226
Qy      60 LeuSerLysLysThrGlyProValPheThrIleTyrLeuGlyProTrpArgProValVal 79
Db      227 CTGAGCAAGAGATGATGCTCCATGTACACAGTGCACCTCTGGACCC---AGCGGGTGTG 283
Qy      80 ValLeuValGlyGlnGlnAlaValArgGlnAlaLeuGlyGlyGlnAlaGluGluPheSer 99
Db      284 GTCTCTCAGCGGTACCAAGCTGTGAAGAGGCGCTGTGGTGGACAGGAGGAGGAGGAGTTAGT 343
Qy      100 GlyArgGlyThrValAlaMetLeuGluGlyThrPheAspGlyHisGlyValPhePheSer 119
Db      344 GGCGCGGCTGACTACCTGCTCTTTTCAACTTTTACCAAGGGCAATGCGATCGCTCTCTCC 403
Qy      120 AsnGlyGluArgTrpArgGlnLeuArgLysPheThrMetLeuAlaLeuArgAspLeuGly 139
Db      404 AGTGGGATCGATGGAAGGTCTCTGAGACAGTTCTCTATCCAGATTCTACGGAATTCGGG 463
Qy      140 MetGlyLysArgGluGlyGluLeuIleGlnAlaAlaArgCysLeuValGluThr 159
Db      464 ATGGGGAAGAGAGCATTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 523
Qy      160 PheGlnGlyThrGluGlyArgProPheAspProSerLeuLeuAlaGlnAlaThrSer 179

```



QY    489 GlnLeu 490  
       |||||  
Db    1462 CAGCTG 1467

RESULT 9  
US-08-238--821B-4  
; Sequence 4, Application US/08238821B  
; Patent No. 591210  
; GENERAL INFORMATION:  
; APPLICANT: GOLDSTEIN, Joyce A.  
; APPLICANT: ROMKES-SPARKS, Marjorie  
; APPLICANT: DE MORAIS, Sonia M.F.  
; TITLE OF INVENTION: CLONING, EXPRESSION AND DIAGNOSIS OF HUMAN  
; TITLE OF INVENTION: CYTOCHROME P450 2C19: THE PRINCIPAL DETERMINANT OF S-  
; TITLE OF INVENTION: MEPHENYTOIN METABOLISM  
; NUMBER OF SEQUENCES: 61  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, 8th Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: US  
; ZIP: 94111  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/239,821B  
; FILING DATE: 06-MAY-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/201,118  
; FILING DATE: 22-FEB-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/864,962  
; FILING DATE: 09-APR-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Liebeschuetz, Joe  
; REGISTRATION NUMBER: 37,505  
; REFERENCE/DOCKET NUMBER: 15280-192110US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (650) 326-2400  
; TELEFAX: (650) 326-2422  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1854 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; FEATURE:  
; NAME/KEY: Region  
; LOCATION: 1..12  
; OTHER INFORMATION: /note= "Corresponds to positions -12 to-1  
; OTHER INFORMATION: for 25 of Figure 2."  
  
US-08-238-821B-4

Alignment Scores:
Pred. No.: 1.31e+108 Length: 1854
Score: 1087.50 Matches: 216
Percent Similarity: 65.35% Conservative: 99
Best Local Similarity: 44.81% Mismatches: 164
Query Match: 41.59% Indels: 3
DB: 2 Gaps: 3

US-10-669-693-2 (1-504) x US-08-238-821B-4 (1-1854)

QY    9 LeuLeuLeuAlaLeuAlaLeuLeuLeuThrLeuAlaLeuSerGlyThrArgAla 28  
Db    31 CTTGTGCCTCTGTCATGGTTCCTCTCTCTTCACTCGGACAGAGCTCT---GGG 87

QY 129 LysPheThrMetLeuAlaLeuArgAspLeuGlyMetGlyLysArgGluGlyGluLeu 148  
Db 385 CGTTTCTCCCTCATGACGCTGCGGAATTTTGGATGGGAAGAGGACATTGAGGACCGT 444  
QY 149 IleGlnAlaGluAlaArgCysLeuValGluThrPheGlnGlyThrGluGlyArgProPhe 168  
Db 445 GTTCAAGAGGAAGCCCGCTGCTTGTGGAGGAGTTTGAGAAAACCAAGCCCTCACCTGT 504  
QY 169 AppProSerLeuLeuLeuAlaGlnAlaThrSerAsnValValCysSerLeuLeuPheGly 188  
Db 505 GATCCCACTTTCATCTCGGGCTGTGCTCCCTGCATGTGATCTGCTCCATTTATTTTCAT 564  
QY 189 LeuArgPheSerTyrGluAspLysGluPheGlnAlaValValArgAlaAlaGlyGlyThr 208  
Db 565 AAACGTTTTTCATTATAAAGATCAGCAATTTCTTAACCTAATGCAAAAGTTGAATGAAAAC 624  
QY 209 LeuLeuGlyValSerSerGlnGlyGlnThrTyrGluMetPheSerTyrPheLeuArg 228  
Db 625 ATCAAGATTTTGAAGCAGCCCTGGATCCAGATCTGCAATAATTTTCTCTCATCATGAT 684  
QY 229 ProLeuProGlyProHisLysGlnLeuLeuHisValSerThrLeuAlaAlaPheThr 248  
Db 685 TACTTCCGGGAACCTACACAATTTACTTAAAAACGTTGCTTTTATGAAAAGTTATATT 744  
QY 249 ValArgGlnValGlnGlnHisGlnGlyAsnLeuAspAlaSerClyProAlaArgAspLeu 268  
Db 745 TTGAAAAAGTAAAAAGAACCAAGAAATCAATGGACATGAACAACCT---CAGACTTT 801  
QY 269 ValAspAlaPheLeuLeuLysMetAlaGlnGluGlnAsnProGlyThrGluPheThr 288  
Db 802 ATTGATTGCTTCTCGTGAATAATGGAGAGGAAGAAACACACCAACCATCTGAATTACT 861  
QY 289 AsnLysAsnMetLeuMetThrValIleTyrLeuLeuPheAlaGlyThrMetThrValSer 308  
Db 862 ATTGAAAGCTTGGAAAAACACTGCAGTTGACTTGTGTGGAGCTGGGACAGACGACAAG 921  
QY 309 ThrThrValGlyTyrThrLeuLeuLeuMetLysTyrProHisValGlnLysTrpVal 328  
Db 922 ACAACCTGAGATATGCT 981  
QY 329 ArgGluGluLeuAsnArgGluLeuGlyValaGlyValaGlnAlaProSerLeuGlyAspArgThr 348  
Db 982 CAGAAAGAGATTGAACGTGTGATTGGCAGAAAACCGGAGCCCTGTCATGCAAGACAGAG 1041  
QY 349 ArgLeuProTyrThrAspAlaValLeuHisGluAlaGlnArgLeuLeuAlaLeuValPro 368  
Db 1042 CACATGCCTACACAGATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1101  
QY 369 MetGlyIleProArgThrLeuMetArgThrThrArgPheArgGlyTyrThrLeuProGln 388  
Db 1102 ACCAGCCTGCCTCCATGTCAGTGTGACCTGTGACATTAATTCAGAAACTATCTCATCC 1161  
QY 389 GlyThrGluValPheProLeuLeuGlyLysIleLeuHisAspProAsnIlePheLysHis 408  
Db 1162 GGCAACAACCATTAATTTCTCTGACTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1221  
QY 409 ProGluGluPheAsnProAspArgPheLeuAspAlaAspGlyArgPheArgLysHisGlu 428  
Db 1222 CCAGAGATGTTTGACCTTCATCCTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1281  
QY 429 AlaPheLeuProPheSerLeuGlyLysArgValCysLeuGlyGluLeuAlaLysAla 448  
Db 1282 TACTTCATGCCCTTCTCAGCAGGAAAAACGGATTTGTGTGGAGAACCTTGGCCGCGCATG 1341  
QY 449 GluLeuPheLeuPhePheThrThrIleLeuGlnAlaPheSerLeuGluSerProCysPro 468  
Db 1342 GAGCTGTGTTTTTATTCCTGACCTCCATTTTACAGAACTTTAACCTGGAATCTCTGTGTG 1401  
QY 469 ProAspThrLeuSerLeuLysProThrValSerGlyLeuPheAsnIleProProAlaPhe 488  
Db 1402 CCAAGAGACCTTGACACCACTCCAGTTGTCATGTGTTGCTCTGTGTGCGCCCTTCTAC 1461

QY 29 ArgGlyHisLeuProGlyProThrProLeuLeuGlyAsnLeuLeuGlnLeu 48  
DB 88 AGAGGAAACTCCCTCTCGGCCCTCTCTCTCCAGTGAATTCCTACAGATA 147  
QY 49 ArgProGlyAlaLeuTyrSerGlyLeuMetArgLeuSerTyrGlyProValPhe 68  
DB 148 GGTATTAAAGCAGCAGCAAACTCTTAAACCAATCTCTCAAGGCTCTATGGCCCTGTGTC 207  
QY 69 ThrIleTyrLeuGlyProThrArgProValValLeuValGlnGlnAlaValArg 88  
DB 208 ACTCTGTATTGGC---CTGAACCACTAGTGTCTCATGATATGAAGCAGTGAAG 264  
QY 89 GluAlaLeuGlyGlyGlnAlaGluGluPheSerGlyValGlyThrValAlaMetLeuGlu 108  
DB 265 GAAAGCCTGATTGATCTTGGAGAGGAGTCTTCTGGAAGGAGCATTTCCCATCGGCTGAA 324  
QY 109 GlyThrPheAspGlyHisGlyValPhePheSerAsnGlyGluArgTrpArgGlnLeuArg 128  
DB 325 AGAGCTAACAGAGGATTTGGAAATGTTTTCAGCAATGGAAGAAATGGAAGGAGATCCGG 384  
QY 129 LysPheThrMetLeuAlaLeuArgAspLeuGlyMetGlyValysArgGluGlyGlnLeu 148  
DB 385 CGTTCTCCCTCATGACGCTGCGGAATTTTGGATGGGAGAGGAGCATTCAGGACCGT 444  
QY 149 IleGlnAlaGluAlaArgCysLeuValGluThrPheGlnGlyThrGluGlyArgProPhe 168  
DB 445 GTTCAAGAGGAAGCCGCTGCTTGGAGGAGTTGAGAAAACCAAGCCCTCACCTGT 504  
QY 169 AspProSerLeuLeuAlaGlnAlaThrSerAsnValValCysSerLeuLeuPheGly 188  
DB 505 GATCCCATTTCACTCTCGGCTGTCTCTCTGCAATGTGATCTGCTCCATTATTTCAT 564  
QY 189 LeuArgPheSerTyrGluAspLysGluPheGlnAlaValAlaArgAlaGlyGlyThr 208  
DB 565 AAACGTTTGTATTAAAGATCAGCAATTTCTTAACCTTAATGAGAAAGTTGAATGAAAC 624  
QY 209 LeuLeuGlyValSerSerGlnGlyGlnThrTyrGluMetPheSerTrpPheLeuArg 228  
DB 625 ATCAAGATTTTGAGCAGCCCTGGATCCAGATCTGCAATATTTTCTCTCATCATGAT 684  
QY 229 ProLeuProGlyProHisLysGlnLeuLeuHisValSerThrLeuAlaAlaPheThr 248  
DB 685 TACTTCCCGGAATCAACAAATTAATCTTAAACGTTGCTTTTGAAGAAAGTTATATT 744  
QY 249 ValArgGlnValGlnGlnHisGlnGlnAsnLeuAspAlaSerGlyProAlaArgAspLeu 268  
DB 745 TTGAAAAGTAAAGAACCAACCAAGATCAATGGACATGAACAAACCTT---CAGGACTTT 801  
QY 269 ValAspAlaPheLeuLeuLysMetAlaGlnGluGlnAsnProGlyThrGluPheThr 288  
DB 802 ATTGATTGCTTCTGTGATAATGGAGAGGAAAGCAGCAACCAACCATCTGAATTTACT 861  
QY 289 AsnLysAsnMetLeuMetThrValIleTyrLeuLeuPheAlaGlyThrMetThrValSer 308  
DB 862 ATTTAAAGCTTGGAAAACATCGCTGACTTGTGGAGCTGGGACAGACAGCAAGC 921  
QY 309 ThrThrValGlyTyrThrLeuLeuLeuMetLysTyrProHisValGlnLysTrpVal 328  
DB 922 ACAACCTGAGATATGCT 981  
QY 329 ArgGluGluLeuAsnArgGluLeuGlyValAlaGlnAlaProSerLeuGlyAspArgThr 348  
DB 982 CAGGAAGAGATTGAACGTGTGATTGGCAGAAACCGGAGCCCTGCATGCAAGCAGGAG 1041  
QY 349 ArgLeuProTyrThrAspAlaValLeuHisGlnAlaGlnArgLeuLeuAlaLeuValPro 368  
DB 1042 CACATGCCCTACAGATGCTGTGTGTGACAGAGTCCAGAGATACCTTGACCTTCTCCCC 1101  
QY 369 MetGlyIleProArgThrLeuMetArgThrThrArgPheArgGlyTyrThrLeuProGln 388  
DB 1102 ACCAGCCTGCCCATGACGTGACCTGTGACATTAATTCAGAAACTATCTCTATCCCAAG 1161

QY 389 GlyThrGluValPheProLeuLeuGlySerIleLeuHisAspProAsnIlePheLysHis 408  
DB 1162 GGCAACAACATATAATTTTCCCTGACTTCTGTCTACATGACAAAGAAATTTCCCAAC 1221  
QY 409 ProGluGluPheAsnProAspArgPheLeuAspAlaAspGlyArgPheArgLysHisGlu 428  
DB 1222 CCAGAGATGTTTGGACCTCATCACTTTCTGGATGAAGGTGGCAATTTTAAAGAAAGTAAA 1281  
QY 429 AlaPheLeuProPheSerLeuGlyLysArgValCysLeuGlyGluGlyLeuAlaLysAla 448  
DB 1282 TACTTTCATGCCITTTCTCAGCAGGAAACCGATTTGTGGGAGAACCCCTGGCCGCATG 1341  
QY 449 GluLeuPheLeuPhePheThrThrIleLeuGlnAlaPheSerLeuGluSerProCysPro 468  
DB 1342 GAGCTGTGTTTTTATCTGACCTCCATTTTACAGAACTTTTAACTGAAATCTCTGTGTTGAC 1401  
QY 469 ProAspThrLeuSerLeuLysProThrValSerGlyLeuPheAsnIleProAlaPhe 488  
DB 1402 CCAAGAACTTGACACCACTCCAGTTGCAATGTTTTCCTCTGTGCGCCCTTCTAC 1461  
QY 489 GlnLeu 490  
DB 1462 CAGCTG 1467

RESULT 10  
US-09-023-655-1056  
; Sequence 1056, Application US/09023655  
; Patent No. 6607879  
; GENERAL INFORMATION:  
; APPLICANT: Cocks, Benjamin G.  
; APPLICANT: Susan G. Stuart  
; APPLICANT: Jeffrey J. Seilhamer  
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE  
; TITLE OF INVENTION: EXPRESSION  
; NUMBER OF SEQUENCES: 1508  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
; STREET: 3174 PORTER DRIVE  
; CITY: PALO ALTO  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/023,655  
; FILING DATE: HEREWITH  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Zeller, Karen J.  
; REGISTRATION NUMBER: 37,071  
; REFERENCE/DOCKET NUMBER: PA-0001 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (650) 855-0555  
; TELEFAX: (650) 845-4166  
; INFORMATION FOR SEQ ID NO: 1056:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1854 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: GENBANK  
; CLONE: g181301  
US-09-023-655-1056















Search completed: September 18, 2004, 04:35:53  
Job time : 153 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 18, 2004, 03:29:12 ; Search time 732 Seconds

(without alignments)  
3475.291 Million cell updates/sec

Title: US-10-669-693-2

Perfect score: 2615

Sequence: 1 MEATGTWALLALALLLLT.....PPAFQLQVRVTLHSTHTQR 504

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 3327077 seqs, 2523723180 residues

Total number of hits satisfying chosen parameters: 6654154

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+\_p2n.model -DEV=xlh

-Q=/cgn2\_1/USPTO.spool/US10669693/runat\_15092004\_101312\_5958/app\_query.fasta\_1.647

-DB=Published Applications NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1

-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62

-TRANS=human40.cdi -LIST=45 -LOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100

-THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZB=500 -MINLEN=0

-MAXLEN=2000000000 -USER=US10669693@cgn\_1\_1\_533@runat\_15092004\_101312\_5958

-NCPUS=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100

-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=120 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5

-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:

1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:

2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:

3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:

4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:

5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:

6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:

7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:

8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:

9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq:

10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq:

11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:

12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:

13: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq:

14: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq:

15: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq:

16: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq:

17: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:

18: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:

19: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Match	Length	ID	Description
------------	-------	-------	--------	----	-------------

1	2615	100.0	1515	15	US-10-067-668-6	Sequence 6, Appli
2	2615	100.0	1515	15	US-10-175-696-6	Sequence 6, Appli
3	2615	100.0	1515	16	US-10-231-913-39	Sequence 39, Appli
4	2615	100.0	1515	17	US-10-776-871-6	Sequence 6, Appli
5	2615	100.0	1927	15	US-10-067-668-4	Sequence 4, Appli
6	2615	100.0	1927	15	US-10-175-696-4	Sequence 4, Appli
7	2615	100.0	1927	17	US-10-776-871-4	Sequence 4, Appli
8	2615	100.0	2085	16	US-10-120-988-436	Sequence 436, App
9	2615	100.0	2604	17	US-10-363-829-102	Sequence 102, App
10	2615	100.0	2620	13	US-10-258-080-16	Sequence 16, Appl
11	2615	100.0	2623	13	US-10-147-493-63	Sequence 63, Appl
12	2615	100.0	2623	13	US-10-145-127-63	Sequence 63, Appl
13	2615	100.0	2623	13	US-10-160-503-63	Sequence 63, Appl
14	2615	100.0	2623	13	US-10-143-118-63	Sequence 63, Appl
15	2615	100.0	2623	13	US-10-144-993-63	Sequence 63, Appl
16	2615	100.0	2623	13	US-10-158-787-63	Sequence 63, Appl
17	2615	100.0	2623	13	US-10-140-024-63	Sequence 63, Appl
18	2615	100.0	2623	13	US-10-140-808-63	Sequence 63, Appl
19	2615	100.0	2623	13	US-10-152-405-63	Sequence 63, Appl
20	2615	100.0	2623	13	US-10-127-852A-63	Sequence 63, Appl
21	2615	100.0	2623	13	US-10-127-900A-63	Sequence 63, Appl
22	2615	100.0	2623	13	US-10-128-685A-63	Sequence 63, Appl
23	2615	100.0	2623	13	US-10-131-820A-63	Sequence 63, Appl
24	2615	100.0	2623	13	US-10-142-886-63	Sequence 63, Appl
25	2615	100.0	2623	13	US-10-146-728-63	Sequence 63, Appl
26	2615	100.0	2623	13	US-10-146-786-63	Sequence 63, Appl
27	2615	100.0	2623	13	US-10-147-499-63	Sequence 63, Appl
28	2615	100.0	2623	13	US-10-157-798-63	Sequence 63, Appl
29	2615	100.0	2623	15	US-10-028-072-63	Sequence 63, Appl
30	2615	100.0	2623	15	US-10-121-049-63	Sequence 63, Appl
31	2615	100.0	2623	15	US-10-123-904-63	Sequence 63, Appl
32	2615	100.0	2623	15	US-10-140-470-63	Sequence 63, Appl
33	2615	100.0	2623	15	US-10-175-746-63	Sequence 63, Appl
34	2615	100.0	2623	15	US-10-176-918-63	Sequence 63, Appl
35	2615	100.0	2623	15	US-10-176-921-63	Sequence 63, Appl
36	2615	100.0	2623	15	US-10-137-865-63	Sequence 63, Appl
37	2615	100.0	2623	15	US-10-140-474-63	Sequence 63, Appl
38	2615	100.0	2623	15	US-10-142-431-63	Sequence 63, Appl
39	2615	100.0	2623	15	US-10-143-114-63	Sequence 63, Appl
40	2615	100.0	2623	15	US-10-140-002-63	Sequence 63, Appl
41	2615	100.0	2623	15	US-10-142-419-63	Sequence 63, Appl
42	2615	100.0	2623	15	US-10-123-262-63	Sequence 63, Appl
43	2615	100.0	2623	15	US-10-142-423-63	Sequence 63, Appl
44	2615	100.0	2623	15	US-10-121-050-63	Sequence 63, Appl
45	2615	100.0	2623	15	US-10-141-755-63	Sequence 63, Appl

#### ALIGNMENTS

##### RESULT 1

US-10-067-668-6

; Sequence 6, Application US/10067668

; Publication No. US20030022334A1

; GENERAL INFORMATION:

; APPLICANT: Glucksmann, Maria Alexandra

; TITLE OF INVENTION: 33303, 32579, NOVEL HUMAN

; TITLE OF INVENTION: CYTOCHROME P450 FAMILY MEMBERS AND USES THEREOF

; FILE REFERENCE: 10448-136001

; CURRENT APPLICATION NUMBER: US/10/067,668

; CURRENT FILING DATE: 2002-02-04

; PRIOR APPLICATION NUMBER: 60/266,140

; PRIOR FILING DATE: 2001-02-02

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 6

; LENGTH: 1515

; TYPE: DNA

; ORGANISM: Homo sapiens

; US-10-067-668-6

Result

Query

Match

Length

ID

Description

Alignment Scores:

Pred. No.: 2.52e-314

Length:

1515

Score:	2615.00	Matches:	504
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	15	Gaps:	0
US-10-669-693-2 (1-504) x US-10-067-668-6 (1-1515)			
QY	1	MetGluAlaThrGlyThrTrpAlaLeuLeuLeuAlaLeuAlaLeuLeuLeuLeuLeuThr	20
DB	1	ATGGAGGCGACCGGCACCTTGCTGGCGCTGCTGGCGCTGCTCTGCTGCTGCTGAGG	60
QY	21	LeuAlaLeuSerGlyThrArgAlaArgGlyHisLeuProGlyProGlyProThrProLeuPro	40
DB	61	CTGGCGCTGTCCGGGACCAGGGCCCGAGCGCACCTTGCCCCCGGGCCACGCGCGTACCA	120
QY	41	LeuLeuGlyAsnLeuLeuGlnLeuArgProGlyAlaLeuTyrSerGlyLeuMetArgLeu	60
DB	121	CTGCTGGGAAACCTCTGCAGCTAGCGCCCGGGGCGCTGATTTCAGGGCTCATGGGGTG	180
QY	61	SerLysLysTyrGlyProValPheThrIleTyrLeuGlyProTrpArgProValValVal	80
DB	181	AGTAAGAAGTACGGACCGGTGTTCACCATCTACCTGGGACCGTGGCGGCTGTGGTGTC	240
QY	81	LeuValGlyGlnGluAlaValArgGluAlaLeuGlyGlyGlnAlaGluPheSerGly	100
DB	241	CTGGTTGGCGAGGCGTGTGCGGAGGCCTTGGGAGGTCAGGCTCAGGAGTTCAGCGGC	300
QY	101	ArgGlyThrValAlaMetLeuGluGlyThrPheAspGlyHisGlyValPhePheSerAsn	120
DB	301	CGGGGAACGTTAGCATGCTGGAAGGAGCTTTGATGGCCATGGGGGTTTCTTCTCCAAAC	360
QY	121	GlyGluArgTrpArgGlnLeuArgLysPheThrMetLeuAlaLeuArgAspLeuGlyMet	140
DB	361	GGGAGCGGTGGAGCGAGCTGAGGAAGTTTACATGCTTGCTCTGCGGGACCTGGCGATG	420
QY	141	GlyLysArgGluGlyGluLeuLeuGlnAlaGluAlaArgCysLeuValGluThrPhe	160
DB	421	CGGAAAGCGAAGGCGAGGAGCTGATCCAGGCGGAGGCCCGGTGTCTGTGTGGAGACATTC	480
QY	161	GlnGlyThrGluGlyArgProPheAspProSerLeuLeuLeuAlaGlnAlaThrSerAsn	180
DB	481	CAGGGGACAGAAGGAGCCCATTCGATCCCTCTGCTGCTGGCCAGGCGACCTCCAAAC	540
QY	181	ValValCysSerLeuLeuPheGlyLeuArgPheSerTyrGluAspLysGluPheGlnAla	200
DB	541	GTAGTCTGCTCCCTCTCTTTGGCCTCCGCTTCTCTATGAGATAAGGAGTTCAGGGCC	600
QY	201	ValValArgAlaAlaGlyGlyThrLeuLeuGlyValSerSerGlnGlyGlyGlnThrTyr	220
DB	601	GTGGTCCGGGACGCTGGTGGTACCCCTGCTGGGAGTCAGCTCCAGGGGGGTTCAGACTAC	660
QY	221	GluMetPheSerTrpPheLeuArgProLeuProGlyProHisLysGlnLeuLeuHis	240
DB	661	GAGATGTTCTCTCGTCTCTGCGGCCCTTCGCCAGGCCCCCAACAGCAGCTCCTCCACCAC	720
QY	241	ValSerThrLeuAlaAlaPheThrValArgGlnValGlnGlnHisGlnGlyAsnLeuAsp	260
DB	721	GTCCAGCACCTTGCGCTGCCCTTCACAGTCCCGCGAGGTTCAGCAGCACCCAGGGGAACCTGGAT	780
QY	261	AlaSerGlyProAlaArgAspLeuValAspAlaPheLeuLeuLysMetAlaGlnGluGlu	280
DB	781	GTTTCGGGCCCGCACGTGACCTTGTGATGCCTTCTGCTGNAAGATGGCAGAGAGGAA	840
QY	281	GlnAsnProGlyThrGluPheThrAsnLysAsnMetLeuMetThrValIleTyrLeuLeu	300
DB	841	CAAAACCCAGGCACAGAAATTCACCAACAAGAACATGCTGATGACAGTCAATTTATTGCTG	900
QY	301	PheAlaGlyThrMetThrValSerThrThrValGlyTyrThrLeuLeuLeuMetLys	320
DB	901	TTTGCTGGGACGATGACGGTCAGCACCCAGGTCGGCTATACCTCTCTGCTCTCTGATGAA	960
QY	321	TyrProHisValGlnLysTrpValArgGluGluLeuAsnArgGluLeuGlyAlaGlyGln	340

```

; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1515
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-175-696-6

```

Alignment Scores:

Pred. No.:	2.52e-314	Length:	1515
Score:	2615.00%	Matches:	504
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	15	Gaps:	0

US-10-669-693-2 (1-504) x US-10-175-696-6 (1-1515)

1	Met	Glul	Ala	Thr	Gly	Thr	Trp	Ala	Leu	Leu	Leu	Ala	Leu	Leu	Leu	Leu	Leu	Leu	Thr	20
	QY																			
1	ATG	GAG	GCG	GAC	CGG	CAC	CTG	GCG	GCT	GCG	GCT	GCG	GCT	GCG	GCT	GCG	GCT	GCG	GCT	GA
	Db																			CG
21	Leu	Ala	Leu	Ser	Gly	Thr	Arg	Ala	Arg	Gly	His	Leu	Pro	Gly	Pro	Gly	Pro	Thr	Pro	Leu
	QY																			Pro
61	CTG	GCG	GCT	GCT	CGG	GAC	CAG	GCG	CCG	AGG	CC	ACT	GCG	CCG	CCG	CCG	CCG	CCG	CCG	TAC
	Db																			CA
41	Leu	Leu	Gly	Asn	Leu	Leu	Gln	Leu	Arg	Pro	Gly	Ala	Leu	Tyr	Ser	Gly	Leu	Met	Arg	Leu
	QY																			Leu
121	CTG	CTG	GCG	AAA	AA	CT	CCT	CAG	CT	AGG	CCG	CGG	GCG	GCT	GT	TAT	T	CAG	GCG	CT
	Db																			AT
61	Ser	Lys	Tyr	Gly	Pro	Val	Phe	Thr	Ile	Tyr	Leu	Gly	Pro	Trp	Arg	Pro	Val	Val	Val	80
	QY																			
181	AGT	AAG	AAG	TAC	GAC	CGG	TGT	T	CAC	CAT	TAC	CTG	GGA	CCG	TGG	GCG	GCG	CTG	TGG	TGG
	Db																			TC
81	Leu	Val	Gly	Gln	Gl	uAla	Val	Arg	Gl	uAla	Leu	Gly	Gln	Ala	Gl	uGlu	Phe	Ser	Gly	100
	QY																			
241	CTG	GTT	GCG	CAG	GAG	GCT	GTC	GCG	GAG	CCC	CTG	GAG	GCT	CAG	GCT	GAG	GAG	TTC	CAG	GCG
	Db																			GC
101	Arg	Gly	Thr	Val	Ala	Met	Leu	Glu	Gly	Thr	Phe	Asp	Gly	His	sGly	Val	Phe	Phe	Ser	Asn
	QY																			120
301	CGG	GAA	CCG	TAG	CGA	TGCT	GTA	AGG	AAC	TTC	T	GAT	GCC	AT	GGG	GTT	TCT	TCT	CTC	CAAC
	Db																			360
121	Gly	Gl	uArg	T	Arg	Gln	Leu	Arg	Lys	Phe	Thr	Met	Leu	Ala	Leu	Arg	Asp	Leu	Gly	Met
	QY																			140
361	GCG	GAG	CGG	TGG	GCG	AGCT	GAG	GAA	GTT	T	AC	AT	GTT	GCT	CTG	GCG	GAC	CTG	GCG	CA
	Db																			TG
141	Gly	Lys	Arg	Glu	Gly	Glu	Gl	uLeu	Ile	Gln	Ala	Leu	Ala	Arg	Cys	Leu	Val	Gl	uThr	Phe
	QY																			160
421	GGA	GAG	CGA	AAG	CGC	GAG	AGCT	GAT	CAT	CGG	GCG	GAG	CCC	GCG	TGT	CTG	TGG	GAG	CAT	TTC
	Db																			480
161	Gln	Gly	Thr	Gl	uGly	Arg	Pro	Phe	Asp	Pro	Ser	Leu	Leu	Ala	Gln	Ala	Thr	Ser	Asn	180
	QY																			
481	CAG	GGA	CAG	AAG	GAC	CGCC	AT	TGC	AT	CCCT	TCC	CTG	CTG	GCG						

Qy	281	GlnAsnProGlyThrGluPheThrAsnLysAsnMetLeuMetThrValIleTyrLeuLeu	300
Db	841	CAAAACCCAGGCGACAGAAATTCACCAACAGAAACATGCTGATGACAGTCATTTATTTGCTG	900
Qy	301	PheAlaGlyThrMetThrValSerThrThrValGlyTyrThrLeuLeuLeuMetLys	320
Db	901	TTTCTGTGGACGATGACGGTCAGCACCAACGGTCGGCTATACCTCTCTGCTCTGATGAAA	960
Qy	321	TyrProHisValGlnLysTrpValArgGluGluLeuAsnArgGluLeuGlyAlaGlyGln	340
Db	961	TACCTCATGTCCAAAAGTGGGTACGTGAGGAGCTGAATCGGAGCTGGGGCTGGCCAG	1020
Qy	341	AlaProSerLeuGlyAspArgThrArgLeuProTyrThrAspAlaValLeuHisGluAla	360
Db	1021	GCACCAAGCCTAGGGGACCGTACCGCGCTCCCTTACACCGACGCGTTCGTCATGAGCGC	1080
Qy	361	GlnArgLeuLeuAlaLeuValProMetGlyIleProArgThrLeuMetArgThrThrArg	380
Db	1081	CAGCGGCTGTGGCGCTGGTGCCCATGGGAATACCCCGACCCCTCATGGGACCAACCCGC	1140
Qy	381	PheArgGlyTyrThrLeuProGlnGlyThrGluValPheProLeuLeuGlySerIleLeu	400
Db	1141	TTCCGAGGGTACACCTCGCCCGACGGGACGGAGGTCTTCCCTCTCTTGGCTCCATCCTG	1200
Qy	401	HisAspProsnIlePheLysHisProGluGluPheAsnProAspArgPheLeuAspAla	420
Db	1201	CATGACCCCAACATCTTCAAGACACCCAGAGAGTTCAACCCAGACCGTTTCTCGATGCA	1260
Qy	421	AspGlyArgPheArgLysHisGluAlaPheLeuProPheSerLeuGlyLysArgValCys	440
Db	1261	GATGACGGTTCAGGAAGCATGAGCGTTCCTGCCCTTCTCTTAGGGAAGCGTGTCTGC	1320
Qy	441	LeuGlyGluGlyLeuAlaLysAlaGluLeuPheLeuPheThrIleLeuGlnAla	460
Db	1321	CTTGGAGAGGCGCTGGCAAAAGCGGAGCTCTTCTCTTCTTCCACCATCTTACAAGCC	1380
Qy	461	PheSerLeuGlnSerProCysProProAspThrLeuSerLeuLysProThrValSerGly	480
Db	1381	TTCTCCCTCGAGAGCGCGTCCCGCGGACACCTCGAGCTCAAGCCACCGTCAGTGGC	1440
Qy	481	LeuPheAsnIleProProAlaPheGlnLeuGlnValArgProThrAspLeuHisSerThr	500
Db	1441	CTTTTCAACATTCCTCCAGGCTTCCAGCTGCAAGTCCGTCCCATGACCTTACTCCACC	1500
Qy	501	ThrGlnThrArg 504	
Db	1501	ACGCAGACCAGA 1512	
RESULT 3			
US-10-231-913-39			
; Sequence 39, Application US/10231913			
; Publication NO. US20040005576A1			
; GENERAL INFORMATION:			
; APPLICANT: Guo, Xiaojia S.			
; APPLICANT: Li, Li			
; APPLICANT: Patturajan, Meera			
; APPLICANT: Shimkets, Richard A.			
; APPLICANT: Casman, Stacie J.			
; APPLICANT: Malyankar, Uriel M.			
; APPLICANT: Tchernev, Velizar T.			
; APPLICANT: Vernet, Corine A.			
; APPLICANT: Spytek, Kimberly A.			
; APPLICANT: Shenoy, Suresh G.			
; APPLICANT: Alsobrook II, John P.			
; APPLICANT: Edinger, Schlomit			
; APPLICANT: Peyman, John A.			
; APPLICANT: Stone, David J.			
; APPLICANT: Ellerman, Karen			
; APPLICANT: Gangolli, Esha A.			
; APPLICANT: Boldog, Ference L.			
; APPLICANT: Colman, Steven D.			
; APPLICANT: Eisen, Andrew J.			

```

; APPLICANT: Liu, Xiaohong
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Spaderina, Steven K.
; APPLICANT: Zertusen, Bryan D.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same

```

Alignment Scores:		
Pred. No.:	2,528-314	1515
Score:	2815.00	504
Percent Similarity:	100.00%	Conservative: 0
Best Local Similarity:	100.00%	Mismatches: 0
Query Match:	100.00%	Indels: 0
DB:	16	Gaps: 0











```
Db 1141 TTCCGAGGGTACACCTGCTCCCGAGGACGAGGCTCTTCCCTCCCTTGGCTCCATCTG 1200
Qy 401 HisAspProAsnIlePheLysHisProGluGluPheAsnProAspArgPheLeuAspAla 420
Db 1201 CATGACCCCAACATTTCAAGCACCCAGAAAGATTCAACCCAGACCGTTTCTGATGCA 1260
Qy 421 AspGlyArgPheArgLysHisGluAlaPheLeuProPheSerLeuGlyLysArgValCys 440
Db 1261 CATGACGCGTTTCAGGAAGCATGAGGCGTTCTCTCCCTCTCTTAGGGAGCGGTCTGC 1320
Qy 441 LeuGlyGluGlyLeuAlaLysAlaGluLeuPhePhePheThrThrIleLeuGlnAla 460
Db 1321 CTTGAGAGGGGCTGGCAAAACGGAGCTCTTCTCTTCCACCAACCATCTCAAGCC 1380
Qy 461 PheSerLeuGluSerProCysProAspThrLeuSerLeuLysProThrValSerGly 480
Db 1381 TTTCTCTGGAGAGCGCTGGCCGGCGGACACCTTGAGCGCTCAAGCCACCGTCACTGC 1440
Qy 481 LeuPheAsnIleProProAlaPheGlnLeuGlnValArgProThrAspLeuHisSerThr 500
Db 1441 CTTTTCACATTCCTCCAGCCTTCCAGCTGCAAGTCCGTCCTCCACTGACCTTCACTCC 1500
Qy 501 ThrGlnThrArg 504
Db 1501 ACGCAGACCAGA 1512

RESULT 8
US-10-120-988-436
; Sequence 436, Application US/10120988
; Publication No. US20030219745A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Goodrich, Ryle
; APPLICANT: Liu, Chenghua
; APPLICANT: Ren, Feiyan
; APPLICANT: Wang, Dunrui
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20030219745A1el Nucleic Acids and
; FILE OF INVENTION: Polypeptides
; CURRENT APPLICATION NUMBER: US/10/120,988
; CURRENT FILING DATE: 2002-04-11
; PRIOR APPLICATION NUMBER: 09/774,528
; PRIOR FILING DATE: 2001-01-30
; NUMBER OF SEQ ID NOS: 441
; SOFTWARE: pt FL_genes Version 2.0
; SEQ ID NO 436
; LENGTH: 2085
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (9)..(1523)
US-10-120-988-436

Alignment Scores:
Pred. No.: 4,12e-314 Length: 2085
Score: 2615.00 Matches: 504
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 16 Gaps: 0

US-10-669-693-2 (1-504) x US-10-120-988-436 (1-2085)

Qy 1 MetGluAlaThrGlyThrAlaLeuLeuAlaLeuAlaLeuLeuLeuLeuLeuLeuLeu 20
Db 9 ATGGAGGCGACCGGACCTGGCGGCTGCTGGCGCTGGCGCTGGCGCTGGCGCTGGCGCTG 68
Qy 21 LeuAlaLeuSerGlyThrArgAlaArgGlyHisLeuProProGlyProThrProLeuPro 40
Db 69 CTGGCGCTGTCGGGACCGAGGGCCCGAGGCGACCTGCCCCCGGGCGGCCACCGCGCTAC 128
```

Db 1209 CATGACCCCAACATCTTCAAGCACCCAGAGAGTTCAACCCAGACCGTTTCTCGATGCA 1268  
Qy 421 AspGlyArgPheArgLysHisGluAlaPheLeuProPheSerLeuGlyLysArgValCys 440  
Db 1269 GATGACGGTTTCAAGAGCATGAGGGTTCTTCCCTTCTCTTAGGGAAGCGTGTCTGC 1328  
Qy 441 LeuGlyGluGlyLeuAlaLysAlaGluLeuPhePhePheThrThrIleLeuGlnAla 460  
Db 1329 CTTGAGAGGGCGCTGGCAAAAGCGAGCTCTTCTCTTCTTCCACCACTCTTACAAGCC 1388  
Qy 461 PheSerLeuGluSerProCysProAspThrLeuSerLeuLysProThrValSerGly 480  
Db 1389 TTTCTCCCTGGAGAGCGCGTGGCCGCGACCCCTGAGCTCAAGCCCAACCGTCAGTGC 1448  
Qy 481 LeuPheAsnIleProProAlaPheGlnLeuGlnValArgProThrAspLeuHisSerThr 500  
Db 1449 CTTTTCACATTTCCCCAGCCCTTCCAGCTGCAAGTCCGTCCCACTGACCTTCACTCCACC 1508  
Qy 501 ThrGlnThrArg 504  
Db 1509 ACGCAGACCAGA 1520

RESULT 9

US-10-363-829-102  
; Sequence 102, Application US/10363829  
; Publication No. US20040142331A1  
; GENERAL INFORMATION:  
; APPLICANT: Jackson, Stuart E.; Lincoln, Stephen E.;  
; APPLICANT: Altus, Christina M.; Dufour, Gerard E.;  
; APPLICANT: Chalup, Michael S.; Jackson, Jennifer L.;  
; APPLICANT: Jones, Anissa L.; Yu, Jimmy Y.;  
; APPLICANT: Wright, Rachel J.; Gietzen, Darryl;  
; APPLICANT: Liu, Tommy F.; Yap, Pierre E.;  
; APPLICANT: Dahl, Christopher R.; Momiya, Monika G.;  
; APPLICANT: Bradley, Diana L.; Rohatgi, Sameer D.;  
; APPLICANT: Harris, Bernard; Roseberry Lincoln, Ann M.;  
; APPLICANT: Gerstin, Jr., Edward H.; Peralta, Careyana H.;  
; APPLICANT: David, Marie H.; Panzer, Scott R.;  
; APPLICANT: Flores, Vincent Z.; Daffo, Abel;  
; APPLICANT: Marwaha, Rakesh; Chen, Alice J.;  
; APPLICANT: Chang, Simon C.; Au, Alan P.;  
; APPLICANT: Imman, Rebekah R.  
; TITLE OF INVENTION: MOLECULES FOR DISEASE DETECTION AND TREATMENT  
; FILE REFERENCE: PT-1183 USN  
; CURRENT APPLICATION NUMBER: US/10/363,829  
; CURRENT FILING DATE: 2003-03-05  
; PRIOR APPLICATION NUMBER: PCT/US01/27628  
; PRIOR FILING DATE: 2001-09-05  
; PRIOR APPLICATION NUMBER: US 60/229,751  
; PRIOR FILING DATE: 2000-09-05  
; PRIOR APPLICATION NUMBER: US 60/229,749  
; PRIOR FILING DATE: 2000-09-05  
; PRIOR APPLICATION NUMBER: US 60/229,750  
; PRIOR FILING DATE: 2000-09-05  
; PRIOR APPLICATION NUMBER: US 60/229,747  
; PRIOR FILING DATE: 2000-09-05  
; PRIOR APPLICATION NUMBER: US 60/229,748  
; PRIOR FILING DATE: 2000-09-05  
; PRIOR APPLICATION NUMBER: US 60/230,583  
; PRIOR FILING DATE: 2000-09-05  
; PRIOR APPLICATION NUMBER: US 60/230,517  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR APPLICATION NUMBER: US 60/230,610  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR APPLICATION NUMBER: US 60/230,597  
; PRIOR FILING DATE: 2000-09-06  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 506  
; SOFTWARE: PERL Program  
; SEQ ID NO 102  
; LENGTH: 2604  
; TYPE: DNA

; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; OTHER INFORMATION: Incyte ID No: LG:337818.2.2000SEP08  
US-10-363-829-102

Alignment Scores:  
Pred. No.: 5,82e-314 Length: 2604  
Score: 2615.00 Matches: 504  
Percent Similarity: 100.00% Conservatives: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 17 Gaps: 0

US-10-669-693-2 (1-504) x US-10-363-829-102 (1-2604)

Qy 1 MetGluAlaThrGlyThrTrpAlaLeuLeuAlaLeuLeuLeuLeuLeuLeuThr 20  
Db 40 ATGGAGGCGACCGGACCTGGGCGCTGCTGGCGCTGGCGCTGCTGCTGCTGCTGCG 99  
Qy 21 LeuAlaLeuSerGlyThrArgAlaArgGlyHisLeuProGlyProGlyProThrProLeuPro 40  
Db 100 CTGGCGCTGTCCGGGACCCAGGGCCCGAGGCCACCTGCCCGCCCGGGGCCACCGCGCTACCA 159  
Qy 41 LeuLeuGlyAsnLeuLeuGlnLeuArgProGlyAlaLeuLeuLeuLeuLeuLeuLeu 60  
Db 160 CTGCTGGGAAACCTCTCGAGCTACGGCCCGGGCGCTGTATTTCAGGGCTCATGCGGCTG 219  
Qy 61 SerLysLysTyrglyProValPheThrIleTyrlLeuGlyProThrArgProValValVal 80  
Db 220 AGTAGAGAGTACGGACCGGTGTACCATCTACCTGGGACCCCTGGGCGCTGTGGTGGTC 279  
Qy 81 LeuValGlyGlnGluAlaValArgGluAlaLeuGlyGlyGlnAlaGluGluPheSerGly 100  
Db 280 CTGTTGGGCGAGGAGCTGTGGCGAGGCGCTGGGAGGTTCAGGCTGAGGAGTTACGGCG 339  
Qy 101 ArgGlyThrValAlaMetLeuGluGlyThrPheAspGlyHisGlyValPhePheSerAsn 120  
Db 340 CGGGGAACCTAGCATGCTGGAAAGGACTTTTGTATGGCCATGGGCTTTTCTTCTCCAAC 399  
Qy 121 GlyGluArgTrpArgGlnLeuArgLysPheThrMetLeuAlaLeuArgAspLeuGlyMet 140  
Db 400 GGGGAGCGGTGGAGGAGCTGAGGAAAGTTTACCATCTGCTCTGGCGGACCTGGGGATG 459  
Qy 141 GlyLysArgGluGlyGluLeuLeuLeuGlnAlaGluAlaArgCysLeuValGluThrPhe 160  
Db 460 GGGAGCGAGAGCGGAGGAGCTGATCCAGCGGAGGCGCGCTGTCTGGTGGAGACATTC 519  
Qy 161 GlnGlyThrGluGlyArgProPheAspProSerLeuLeuLeuAlaGlnAlaThrSerAsn 180  
Db 520 CAGGGGACAGAGGAGCGCCCATTCGATCCCTGCTGCTGGCCACAGGCGACCTCCAAC 579  
Qy 181 ValValCysSerLeuLeuPheGlyLeuArgPheSerTyrglyAspLysGluPheGlnAla 200  
Db 580 GTAGTCTGCTCCCTCTCTTTGGCCTCTCTTATGAGGATAAGGAGTTCCAGGCC 639  
Qy 201 ValValArgAlaAlaGlyGlyThrLeuLeuGlyValSerSerGlnGlyGlyGlnThrTyr 220  
Db 640 GTGTCGCGGCGAGCTGTGTGTATCCCTGCTGGAGTCACTCCAGGGGGGTTCAGACCTAC 699  
Qy 221 GluMetPheSerTrpPheLeuArgProLeuProGlyProHisLysGlnLeuLeuHisHis 240  
Db 700 GAGATGTTCTCTGTTCTCTGGGCGCCCTGCCAGGCGCCCAACAGAGCTCTCTCCACAC 759  
Qy 241 ValSerThrLeuAlaAlaPheThrValArgGlnValGlnGlnHisGlnGlyAsnLeuAsp 260  
Db 760 GTCAGCACCTTGGCTGCTTTCACAGTCCGGAGGTGAGCAGCAGCAGGAGAACCTGAT 819  
Qy 261 AlaSerGlyProAlaArgAspLeuValAlaPheLeuLeuLysMetAlaGlnGluGlu 280  
Db 820 GCCTTCGGGCGCGCACCTGACCTTGTGATGCTTCTGCTGAAGATGGCAGAGAGGAA 879  
Qy 281 GlnAsnProGlyThrGluPheThrAsnLysAsnMetLeuMetThrValIleTyrlLeuLeu 300

```
Db 880 CAAACCCAGGACAGATTACCAACAAGAACATGCTGATGACAGTATTATTGCTG 939
Qy 301 PheAlaGlyThrMetThrValSerThrThrValGlyThrThrLeuLeuLeuMetLys 320
Db 940 TTTCTGGAGCATGACGGTCAGACACCGTCGCTATACCTCTCTCTCTGATGAA 999
Qy 321 TyrProHisValGlnLysTyrValArgGluGluLeuAsnArgGluLeuGlyAlaGln 340
Db 1000 TACCTCATGTCCAAAGTGGTACGTGAGAGCTGAATCGGAGCTGGGGCTGGCCAG 1059
Qy 341 AlaProSerLeuGlyAspArgThrArgLeuProTyrThrAspAlaValLeuHisGlnAla 360
Db 1060 GCACCAAGCCTAGGGACCGTACCCTCTTACACCGACGCGTTCATGAGCGG 1119
Qy 361 GlnArgLeuLeuAlaLeuValProMetGlyLeuProArgThrLeuMetArgThrArg 380
Db 1120 CAGGGCTGCTGGCGCTGGTCCCATGGAAATACCCGACCTCATGCGGACACCGC 1179
Qy 381 PheArgGlyTyrThrLeuProGlnGlyThrGluValPheProLeuLeuGlySerIleLeu 400
Db 1180 TTCGAGGGTACACCTGCCCGAGGACGAGGTCCTCCCTCTCTCTGCTCATCTG 1239
Qy 401 HisAspProAsnIlePheLysHisProGluGluPheAsnProAspArgPheLeuAspAla 420
Db 1240 CATGACCCCAACATCTTCAAGACACCCAGAGAGTTCAACCCAGACCGTTCCTGATGCA 1299
Qy 421 AspGlyArgPheArgLysHisGluAlaPheLeuProPheSerLeuGlyLysArgValCys 440
Db 1300 GATGAGCGTTACGAGCATGAGCGCTCTCTGCTCTCTCTCTCTCTCTCTCTCTG 1359
Qy 441 LeuGlyGluGlyLeuAlaLysAlaGluLeuPhePheThrThrIleLeuGlnAla 460
Db 1360 CTGGAGAGGGCTGGCAAGAGGGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1419
Qy 461 PheSerLeuGluSerProCysProAspThrLeuSerLeuLysProThrValSerGly 480
Db 1420 TTTCTCTCTGAGAGCGCTGCCCGCGGACACCTGAGCTCAAGCCCGCTGAGTGGC 1479
Qy 481 LeuPheAsnIleProAlaPheGlnLeuGlnValArgProThrAspLeuHisSerThr 500
Db 1480 CTTTTCAACATTTCCCGAGCTTCCAGCTTCCAGTCCGATCCGCTCCCTCTCTCTC 1539
Qy 501 ThrGlnThrArg 504
Db 1540 ACGCAGACCCAGA 1551

RESULT 10
US-10-258-080-16
; Sequence 16, Application US/10258080
; Publication No. US20040029125A1
; GENERAL INFORMATION:
; APPLICANT: Incyte Genomics, Inc.
; APPLICANT: POLICKY, Jennifer L.
; APPLICANT: HAFALIA, April J.A.
; APPLICANT: BURFORD, Neil
; APPLICANT: RING, Huijun Z.
; APPLICANT: LAL, Preeti G.
; APPLICANT: TRIBOULEY, Catherine M.
; APPLICANT: YAO, Monique G.
; APPLICANT: YUE, Henry
; APPLICANT: TANG, Y. Tom
; APPLICANT: ARVIZU, Chandra S.
; APPLICANT: DAS, Debopriya
; APPLICANT: SANJANWALA, Madhusudan M.
; APPLICANT: GANDHI, Ameena R.
; APPLICANT: REDDY, Roopa M.
; APPLICANT: KHAN, Farrah A.
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: RAMKIMAR, Jayalaxmi
; APPLICANT: GRIFFIN, Jennifer A.
; APPLICANT: AU-YOUNG, Janice K.
; TITLE OF INVENTION: DRUG METABOLIZING ENZYMES
```

```
; FILE REFERENCE: PI-0070 USN
; CURRENT APPLICATION NUMBER: US/10/258,080
; CURRENT FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: US 60/203,509
; PRIOR FILING DATE: 2000-05-11
; PRIOR APPLICATION NUMBER: US 60/202,234
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/200,185
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/198,403
; PRIOR FILING DATE: 2000-04-19
; PRIOR APPLICATION NUMBER: PCT/US01/11869
; PRIOR FILING DATE: 2001-04-32
; PRIOR APPLICATION NUMBER: US 60/197,590
; PRIOR FILING DATE: 2000-04-13
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PERL Program
; SEQ ID NO 16
; LENGTH: 2620
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20040029125A1 1272843CB1
US-10-258-080-16
```

```
Alignment Scores:
Pred. No.: 5,87e-314 Length: 2620
Score: 2615.00 Matches: 504
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13 Gaps: 0
```

US-10-669-693-2 (1-504) x US-10-258-080-16 (1-2620)

```
Qy 1 MetGluAlaThrGlyThrTrpAlaLeuLeuLeuAlaLeuLeuLeuLeuLeuThr 20
Db 41 ATGAGGCGCAGCCGCGACCTGGCGCTGCTGGCGCTGGCGCTGCTGCTGCTGCTG 100
Qy 21 LeuAlaLeuSerGlyThrArgAlaArgGlyHisLeuProGlyProGlyProLeuPro 40
Db 101 CTGGCGCTGTCCGGGACAGGCGCCGAGGCCACCTGCCCCCGGCGCCACCGCGCTACCA 160
Qy 41 LeuLeuGlyAsnLeuLeuGlnLeuArgProGlyAlaLeuTyrSerGlyLeuMetArgLeu 60
Db 161 CTGCTGGGAAACCTCTCTGAGCTACGCGCGCGGCGCTGTATTTCAGGGCTCATGGCGCTG 220
Qy 61 SerLysLysTyrGlyProValPheThrIleTyrLeuGlyProTrpArgProValVal 80
Db 221 AGTAAGAAGTACGAGCCGCGTTCACCATCTACTGGGACCTTGGCGCGCTGTGGTGGTC 280
Qy 81 LeuValGlyGlnGluAlaValArgGluAlaLeuGlyGlyGlnAlaGluGluPheSerGly 100
Db 281 CTGTTGGCAGAGGCTGTGGGAGGCGCTTGGAGGCTGAGGCTGAGGAGTTTTCAGCGGC 340
Qy 101 ArgGlyThrValAlaMetLeuGluGlyThrPheAspGlyHisGlyValPhePheSerAsn 120
Db 341 CGGGGAACCGTAGCGATGCTGGAAAGGACTTTTGTATGGCCATGGGGTTTCTTCTTCAAC 400
Qy 121 GlyGluArgTyrArgGlnLeuArgLysPheThrMetLeuAlaLeuArgAspLeuGlyMet 140
Db 401 GGGGAGCGGTGGAGGACGCTGAGGAAGTTTACCATGCTTGTCTCTGGGACCTGGGCGATG 460
Qy 141 GlyLysArgGluGlyGluLeuIleGlnAlaGluAlaArgCysLeuValGluThrPhe 160
Db 461 GGGAGCGGAGAGGCGGAGAGCTGATCCAGGCGGAGGCGCGGTGTCTGTGGAGACATTC 520
Qy 161 GlnGlyThrGluGlyArgProPheAspProSerLeuLeuLeuAlaGlnAlaThrSerAsn 180
Db 521 CAGGGGACAGAGGAGCGCCCATTCATCCCTCTCTGCTGCTGGCCGAGCGGCTTCTTCAAC 580
Qy 181 ValValCysSerLeuLeuPheGlyLeuArgPheSerTyrGluAspLysGluPheGlnAla 200
```

Db	581	GTAGTCTCTCCCTCTCTTTGGCCCTCGCTTCTCTATGAGTAAAGAGTTCCAGGCC	640
Qy	201	ValValArgAlaAlaGlyGlyThrLeuLeuGlyValSerSerGlnGlyGlyGlnThrTyr	220
Db	641	GTGGTCCGGGAGCTGGTGGTACCTCTGCTGGAGTCAGTCCACGGGGGTCAGACCTAC	700
Qy	221	GluMetPheSerTrpPheLeuArgProLeuProGlyProHisLysGlnLeuLeuHisHis	240
Db	701	GAGATGTTCTCTGCTGCTCTGGGCCCTTGCAGGCCCTGCAGGCCCCCAAGCAGCTCTCCACCAC	760
Qy	241	ValSerThrLeuAlaAlaPheThrValArgGlnValGlnGlnHisGlnGlyAsnLeuLeu	260
Db	761	GTACAGACCTTGGCTTTCAGCTCCGGCGAGGTCAGACACACAGGGGAACCTGGAT	820
Qy	261	AlaSerGlyProAlaArgAspLeuValAspAlaPheLeuLeuLysMetAlaGlnGlu	280
Db	821	GCTTCGGGCCCGCGACGTGACCTTGTGATGCTTCTGCTGAAGATGGCACAGGAGAA	880
Qy	281	GlnAsnProGlyThrGluPheThrAsnLysAsnMetLeuMetThrValIleTyrLeuLeu	300
Db	881	CAAAACCCAGGCACAGAAATTCACCAACAAGAACATGCTGATGACAGTCAATTTATTGCTG	940
Qy	301	PheAlaGlyThrMetThrValSerThrThrValGlyTyrThrLeuLeuLeuMetLys	320
Db	941	TTTGCTGGGACGATGACGCTCAGCACCCAGTCGGCTATACCTCTCTGCTCTGTATGAAA	1000
Qy	321	TyrProHisValGlnLysTrpValArgGluGluLeuAsnArgGluLeuGlyGln	340
Db	1001	TACCCTCATGTCCAAAGTGGGTACGTGAGGAGCTGAATCGGGAGCTGGGGGTGGCCAG	1060
Qy	341	AlaProSerLeuGlyAspArgThrArgLeuProTyrThrAspAlaValLeuHisGluAla	360
Db	1061	GCACAAAGCCTAGGGACCGTACCCGCTCCCTTACACCGACGCGTTCCTGATGAGCG	1120
Qy	361	GlnArgLeuLeuAlaLeuValProMetGlyIleProArgThrLeuMetArgThrArg	380
Db	1121	CAGCGGCTCTGCGCTGGTGCCCATGGGAATACCCCGCACCTCATGCGGACCCCG	1180
Qy	381	PheArgGlyTyrThrLeuProGlnGlyThrGluValPheProLeuLeuGlySerIleLeu	400
Db	1181	TTCCGAGGTATACACCTGCCCGGAGGACGAGGTTCTCCCTCTCTGCTCCATCCTG	1240
Qy	401	HisAspProAsnIlePheLysHisProGluGluPheAsnProAspArgPheLeuAspAla	420
Db	1241	CATGACCCCAACATCTCAAGCACCCAGAGAGTTCAACCCAGACCGTTTCTTGGATGCA	1300
Qy	421	AspGlyArgPheArgLysHisGluAlaPheLeuProPheSerLeuGlyLysArgValCys	440
Db	1301	GATGGACGGTTACAGGAAGCATGAGGGCTTCCTGCCCTTCTCTTAGGGAAGCTGTCTGC	1360
Qy	441	LeuGlyGluGlyLeuAlaLysAlaGluLeuPhePhePheThrThrIleLeuGlnAla	460
Db	1361	CTTTGGAGAGGCTGGCAAAAGCGGAGCTTCTCTCTTTCACCAACCATCTTACAAGCC	1420
Qy	461	PheSerLeuGluSerProCysProProAspThrLeuSerLeuLysProThrValSerGly	480
Db	1421	TTCTCTCTGAGAGCCGTCGCCCGCGACACCTGAGCCTCAAGCCCAACCGTCAGTGGC	1480
Qy	481	LeuPheAsnIleProProAlaPheGlnLeuGlnValArgProThrAspLeuHisSerThr	500
Db	1481	CTTTTCAACATTTCCCCAGCCTTCCAGCTGCAAGTCGTCCTCCACTGACCTTCACTCCACC	1540
Qy	501	ThrGlnThrArg	504
Db	1541	ACGCAGACCAGA	1552

RESULT 11  
US-10-147-493-63  
; Sequence 63, Application US/10147493  
; Publication No. US20040029217A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.

APPLICANT: Beresini, Maureen  
APPLICANT: DeForge, Laura  
APPLICANT: Desnoyers, Luc  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Gerritsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Sherwood, Steven  
APPLICANT: Smith, Victoria  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K  
APPLICANT: Wood, William  
APPLICANT: Zhang, Zemin

TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE REFERENCE: P3330R1C345  
CURRENT APPLICATION NUMBER: US/10/147,493  
CURRENT FILING DATE: 2002-05-17  
Prior Application removed - See File Wrapper or Palm  
NUMBER OF SEQ ID NOS: 550  
SEQ ID NO 63  
LENGTH: 2623  
TYPE: DNA  
ORGANISM: Homo Sapien  
US-10-147-493-63

Alignment Scores:  
Pred. No.: 5,88e-314 Length: 2623  
Score: 2615.00 Matches: 504  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 13 Gaps: 0  
US-10-669-693-2 (1-504) x US-10-147-493-63 (1-2623)

Qy 1 MetGluAlaThrGlyThrTrpAlaLeuLeuLeuAlaLeuAlaLeuLeuLeuLeuThr 20  
Db 41 ATGGAGGCGACCGGACCTGGGCGCTGCTGGCGCTGGCGCTGCTGCTGCTGCTGACG 100

Qy 21 LeuAlaLeuSerGlyThrArgAlaArgGlyHisLeuProProGlyProThrProLeuPro 40  
Db 101 CTGGCGCTGCTCGGGACCGAGGCCCGAGGCCACCTGCCCGGCCCGCCGCGCTACCA 160

Qy 41 LeuLeuGlyAsnLeuLeuGlnLeuArgProGlyAlaLeuTyrSerGlyLeuMetArgLeu 60  
Db 161 CTGCTGGGAAACCTCTCGAGCTACGCGCCCGGGCGCTGTATTTCAGGGCTCATGCGGCTG 220

Qy 61 SerLysLysTyrGlyProValPheThrIleTyrLeuGlyProTrpArgProValValVal 80  
Db 221 AGTAAGAAGTACGAGACCGGTGTTTACCATCTACCTGGGACCTTGGCGGCTGTGGTGGTC 280

Qy 81 LeuValGlyGlnGluAlaValArgGluAlaLeuGlyGlyGlnAlaGluGluPheSerGly 100  
Db 281 CTGGTTGGCAGAGGCTGTGGGAGGCCCTGGGAGGTCAGGCTGAGGAGTTTCAGCGGC 340

Qy 101 ArgGlyThrValAlaMetLeuGluGlyThrPheAspGlyHisGlyValPhePheSerAsn 120  
Db 341 CGGGGAACCGTAGCGATGCTGGAAAGGACTTTTGTATGGCCATGGGTTTTTCTTCTCCAAC 400

Qy 121 GlyGluArgTrpArgGlnLeuArgLysPheThrMetLeuAlaLeuArgAspLeuGlyMet 140  
Db 401 GGGGAGCGGTGGAGGCGAGCTGAGAAAGTTTACCATGCTTGTCTGCGGGACCTGGGCAATG 460

Qy 141 GlyLysArgGluGlyGluGluLeuIleGlnAlaGluAlaArgCysLeuValGluThrPhe 160  
Db 461 GCGAAGCGAAGAGCGGAGGAGCTGATCCAGCGGAGGCCCGGTGCTCTGTGGTGGAGACATTC 520

Qy 161 GlnGlyThrGluGlyArgProPheAspProSerLeuLeuLeuAlaGlnAlaThrSerAsn 180





Qy	161	GlnGlyThrGluGlyArgProPheAspProSerLeuLeuAlaGlnAlaThrSerAsn	180
Db	521	CAGGGACAGAGAGCGCCATTCGATCCCTCCCTGCTGGCCACAGCCACCTCCAAAC	580
Qy	181	ValValCysSerLeuLeuPheGlyLeuAArgPheSerTyrGluAspLysGluPheGlnAla	200
Db	581	GTAGTCTGCTCCCTCTCTTGGCCCTCCGCTTCTCATGAGGATTAAGAGATTCCAGGCC	640
Qy	201	ValValArgAlaAlaGlyGlyThrLeuLeuGlyValSerSerGlnGlyGlyGlnThrTyr	220
Db	641	GTGGTCCGGCAGCTGGTGTGTACCTGCTGGAGTTCAGTCCACAGGGGGGTGAGACTTAC	700
Qy	221	GluMetPheSerTrpPheLeuAArgProLeuProGlyProHisLysGlnLeuLeuHisHis	240
Db	701	GAGATGTTCTCTGGTTCCTCGGCCCTCCAGGCCCCCAACAGCAGCTCTCTCCACCAC	760
Qy	241	ValSerThrLeuAlaAlaPheThrValArgGlnValGlnGlnHisGlnGlyAsnLeuAsp	260
Db	761	GTACGACCTTGGTGTGCTTTCACAGTCCGCGAGTGCAGCAGCACAGGGGAACTGGAT	820
Qy	261	AlaSerGlyProAlaAArgAspLeuValAspAlaPheLeuLeuLysMetalGlnGluGlu	280
Db	821	GCITTCGGGGCCCCGACGTGACCTTGTTCGATGCCCTTCTGCTGAAGATGGCACAGGAGAA	880
Qy	281	GlnAsnProGlyThrGluPheThrAsnLysAsnMetLeuMetThrValLleTyrLeuLeu	300
Db	881	CAAAACCCAGGCA CAGAAATTCACCAACAAGAACATGCTGATGACAGTCAATTTATTTGCTG	940
Qy	301	PheAlaGlyThrMetThrValSerThrThrValGlyTyrThrLeuLeuLeuLeuMetLys	320
Db	941	TTTTCTGGGACGATGAGCGGTGAGCACACCGCTCGGCTATACCTCTCTGCTTCCATGATGAAA	1000
Qy	321	TyrProHisValGlnLysTrpValArgGluGluLeuAsnArgGluLeuGlyValAlaGlyGln	340
Db	1001	TACCTCATGTGTCAAAAGTGGTACGTGAGGAGCTGAATCGGAGAGTGGGGGTGGCCAG	1060
Qy	341	AlaProSerLeuGlyAspArgThrArgLeuProTyrThrAspAlaValLeuHisGluAla	360
Db	1061	GCACCAAGCCTAGGGGACCGTACCCGCTCCCTTTACACCGACGCGGTTCCTGCAATGAGCGG	1120
Qy	361	GlnArgLeuLeuAlaLeuValProMetGlyIleProArgThrLeuMetArgThrThrArg	380
Db	1121	CAGCGGCTGCTGGCGCTGGTGCCCATGGGAAATACCCGCACTCATGCGGACCAACCCGC	1180
Qy	381	PheArgGlyTyrThrLeuProGlnGlyThrGluValPheProLeuLeuGlySerIleLeu	400
Db	1181	TTCCGAGGGTACACCTGCCCCAGGGCACGAGGTCTTCCCTCTCTTGGCTCCATCCTG	1240
Qy	401	HisAspProAsnIlePheLysHisProGluGluPheAsnProAspArgPheLeuAspAla	420
Db	1241	CATGACCCCAACATCTTCAACGACCCAGAGAGTTCAACCAGACCCGTTTCTCGATGCA	1300
Qy	421	AspGlyArgPheArgLysHisGluAlaPheLeuProPheSerLeuGlyLysArgValCys	440
Db	1301	GATGACGGTTCAGGAACATGAGGCGTTCCTGCGCTCTCTTAGGGAAGCGGTCTGCG	1360
Qy	441	LeuGlyGluGlyLeuAlaLysAlaGluLeuPheLeuPhePheThrIleLeuGlnAla	460
Db	1361	CTTGGAGAGGGCTGGCAAAAGCGAGCTCTTCTCTTTCACACCACTCTCTCAAGGCC	1420
Qy	461	PheSerLeuGluSerProCysProPheAspThrLeuSerLeuLysProThrValSerGly	480
Db	1421	TTCTCCCTGGAGAGCCCGTGGCCCGGACACCTTGAGCCTTCAAGCCACCGTCAGTGGC	1480
Qy	481	LeuPheAsnIleProAlaPheGlnLeuGlnValArgProThrAspLeuHisSerThr	500
Db	1481	CTTTTCAACATTTCCCCAGCCCTTCCAGCTGCAAGTCCGTGCCACTGACCTTCACTCCACC	1540
Qy	501	ThrGlnThrArg	504
Db	1541	ACGCAGACAGA	1552

RESULT 13					
US-10-160-503-63					
; Sequence 63, Application US/10160503					
; Publication No. US20040033559A1					
; GENERAL INFORMATION:					
; APPLICANT:	Baker, Kevin P.				
; APPLICANT:	Beresini, Maureen				
; APPLICANT:	DeForge, Laura				
; APPLICANT:	Desnovers, Luc				
; APPLICANT:	Filvaroff, Ellen				
; APPLICANT:	Gao, Wei-Qiang				
; APPLICANT:	Gerritsen, Mary E.				
; APPLICANT:	Goddard, Audrey				
; APPLICANT:	Godowski, Paul J.				
; APPLICANT:	Gurney, Austin L.				
; APPLICANT:	Sherwood, Steven				
; APPLICANT:	Smith, Victoria				
; APPLICANT:	Stewart, Timothy A.				
; APPLICANT:	Tumas, Daniel				
; APPLICANT:	Watanabe, Colin K				
; APPLICANT:	Wood, William				
; APPLICANT:	Zhang, Zemin				
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC ACIDS ENCODING THE SAME					
; FILE REFERENCE: F3330R1C446					
; CURRENT APPLICATION NUMBER: US/10/160,503					
; CURRENT FILING DATE: 2002-05-30					
; NUMBER OF SEQ ID NOS: 550					
; SEQ ID NO 63					
; LENGTH: 2623					
; TYPE: DNA					
; ORGANISM: Homo Sapien					
US-10-160-503-63					
Alignment Scores:					
Pred. No.:	5_88e-314	Length:	2623		
Score:	2615.00	Matches:	504		
Percent Similarity:	100.00%	Conservative:	0		
Best Local Similarity:	100.00%	Mismatches:	0		
Query Match:	100.00%	Indels:	0		
DB:	13	Gaps:	0		
US-10-669-693-2 (1-504) x US-10-160-503-63 (1-2623)					
Qy	1 MetGluAlaThrGlyThrTrpAlaLeuLeuLeuAlaLeuAlaLeuLeuLeuLeuLeuThr	20			
Db	41 ATGGAGCGGCACCGGCCCTGTGTCGTGGCGTGTGTTATCATGGGCTCATGCGTGACG	100			
Qy	21 LeuAlaLeuSerGlyThrArgAlaArgGlyHisLeuProGlyProThrProleupro	40			
Db	101 CTGGCGCTGTCTGGGACCAGGCGCCGAGCCACTTCCCCCGGGCCCACGCCGCTACA	160			
Qy	41 LeuleugLyAsnLeuLeuGlnLeuArgProGlyAlalaLeutyrserysergyleumethargleu	60			
Db	161 CTGCTGGGAACCTCTTCAGCACGCGCCGGCGCTGTATTCAAGGCTCATGCGGTG	220			
Qy	61 SerLysLysTyrglyprovalphethriletyrleuglyprotparqprovalvalVal	80			
Db	221 AGTAAGAAGTAGCGAACCGGTGTTACCATCTACCTGGGACCCCTGGCGGCTGTGTGTGTC	280			
Qy	81 LeuValGlyGlnGlualaValArggluaAlaleuGlyglyGlnAlagluPheSergly	100			
Db	281 CTGGTTGGCAGAGSGCTGTGCGGAGGCCCTTGGAGGTTCAGGCTGAGAGTTTCAGCGGC	340			
Qy	101 ArgGlyThrValalameuleudglythrphesaspGlyHissglyValphePheserAn	120			
Db	341 CGGGAAACCGTAGCCATGTGGAAGGAGCTTTTGATGGCCATGGGTTTTCTTCTCCAAC	400			
Qy	121 GlyGluArgTrpargGlnLeuArglysPheThrMetLeuAlaLeuArgAspleuGlyMet	140			
Db	401 GGGGAGCGGTGGAGCAGCCTGAGGAAGTTTACCATGCTTGCTCTCGGGACCTGGGCA TG	460			
Qy	141 GlyLysArggluglygluLeutleGlnlaGluaAlaArgcysLeuValgluthrPhe	160			

Db	461	GGGAGCGGAGAGCGCAGGAGCTGGATCCAGCGGAGCGCGGTGTCTGGTGGAGACATTC	520
QY	161	GlnGlyThrGluGlyArgProPheAspProSerLeuLeuLeuAlaGlnAlaThrSerAsn	180
Db	521	CAGGGGACAGAGGAGCGCCATTTCGATCCCTCCCTGCTGTCGGCCAGGCGCACCTCCAAAC	580
QY	181	ValValCysSerLeuLeuPheGlyLeuArgPheSerTyrGluAspLysGluPheGlnAla	200
Db	581	GTAGTCTGCTCGCTCTCTTTGGCCTCCGCTTCTCCTATAGGATAAAGAGTTCCAGGGCC	640
QY	201	ValValArgAlaAlaGlyGlyThrLeuLeuGlyValSerSerGlnGlyGlnThrTyr	220
Db	641	GTGGTCCGGGACAGCTGGTGGTATCCCTGCTGGAGTCAGCTCCAGGGGGTTCAGACCTAC	700
QY	221	GluMetPheSerTyrPheLeuArgProLeuProGlyProHisLysGlnLeuHisHis	240
Db	701	GAGATGTTCTCTCTGGTCTCTCGGGCCCTGCGCAGGGCCCCCAAGAAGAGCTCTCTCCACCAC	760
QY	241	ValSerThrLeuAlaAlaPheThrValArgGlnValGlnGlnHisGlnGlyAsnLeuAsp	260
Db	761	GTCAAGACCTTGGCTGCCTTCACAGTCCGGCAGGTGCAGACACACAGGGGACCTGGAT	820
QY	261	AlaSerGlyProAlaArgAspLeuValAspAlaPheLeuLeuLysMetAlaGlnGlu	280
Db	821	GCTTCGGGCGCCGACGTGACCTTCGTGATGCTTCTCTGCTGAAGATGGCACAGGAGAA	880
QY	281	GlnAsnProGlyThrGluPheThrAsnLysAsnMetLeuMetThrValIleTyrLeuLeu	300
Db	881	CAAAACCCAGGCAAGAAATTCACCAACAGAAACATGCTGATGACAGTCAATTTATTTGCTG	940
QY	301	PheAlaGlyThrMetThrValSerThrThrValGlyTyrThrLeuLeuLeuLeuMetLys	320
Db	941	TTTGTGGGACGATGACGGTCAGACACACGCTCGGCTATACCTCCTGCTCTCTGAAGAA	1000
QY	321	TyrProHisValGlnLysTyrValArgGluLeuLeuAsnArgGluLeuGlyAlaGlyGln	340
Db	1001	TACCTCATGTGCCAAAGTGGGTACGTGAGGAGCTGATCGGAGCTGGGGCTGGCCAG	1060
QY	341	AlaProSerLeuGlyAspArgThrArgLeuProTyrThrAspAlaValLeuHisGluAla	360
Db	1061	GCACCAAGCTTAGGGGACCGTACCCTGCTCCCTTACACCGACGGGTTCTCATGAGCGC	1120
QY	361	GlnArgLeuLeuAlaLeuValProMetGlyTyrProArgThrLeuMetArgThrThrArg	380
Db	1121	CAGCGGTGCTGGCGCTGGTCCCATGGGAATACCCCGCACCTCATGCGGACCCCGCCG	1180
QY	381	PheArgGlyTyrThrLeuProGlnGlyThrGluValPheProLeuLeuGlySerIleLeu	400
Db	1181	TTCCGAGGGGTACACCTCGCCCGGACGAGGTCTTCCCTCTCTGGCTCCATCTCTG	1240
QY	401	HisAspProAsnIlePheLysHisProGluGluPheAsnProAspArgPheLeuAspAla	420
Db	1241	CATGACCCCAACATCTTCAAGACACCCAGAGAGTTCAACCCAGACCGCTTCTGGATGCA	1300
QY	421	AspGlyArgPheArgLysHisGluAlaPheLeuProPheSerLeuGlyLysArgValCys	440
Db	1301	GATGGACGGTTCAGGAAGCATGAGCGGTTCTGCCCCCTCTCTTCTTGGGAAGCGTGTCTG	1360
QY	441	LeuGlyGluGlyLeuAlaLysAlaGluLeuPhePheThrThrIleLeuGlnAla	460
Db	1361	CTTGGAGAGGCGCTGGCAAAAGCGGAGCTTCTCTCTTCTTCAACCACTCCCTACAGCC	1420
QY	461	PheSerLeuGlnSerProCysProProAspThrLeuSerLeuLysProThrValSerGly	480
Db	1421	TTCTCCCTGGAGAGCCGTGCGCGCGGACACCTGAGCCTCAAGCCACCGTCAGTGGC	1480
QY	481	LeuPheAsnIleProProAlaPheGlnLeuGlnValArgProThrAspLeuHisSerThr	500
Db	1481	CTTTTCAACATTTCCCGACGCTTCCAGTGCAGTCCGTCCTCCACTGACCTTCACTCCACC	1540
QY	501	ThrGlnThrArg	504

Db 1541 ACGCAGACCAGA 1552

RESULT 14  
US-10-143-118-63  
; Sequence 63, Application US/10143118  
; Publication No. US20040038335A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: DeForge, Laura  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Sherwood, Steven  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: ACIDS ENCODING THE SAME  
; FILE REFERENCE: P333OR1C228  
; CURRENT APPLICATION NUMBER: US/10/143,118  
; CURRENT FILING DATE: 2002-05-09  
; Prior Application removed - See Palm or File Wrapper  
; NUMBER OF SEQ ID NOS: 550  
; SEQ ID NO 63  
; LENGTH: 2623  
; TYPE: DNA  
; ORGANISM: Homo Sapien  
US-10-143-118-63

Alignment Scores:  
Pred. No.: 5,88e-314 Length: 2623  
Score: 2615.00 Matches: 504  
Percent Similarity: 100.00% Conservatative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 13 Gaps: 0

US-10-669-693-2 (1-504) x US-10-143-118-63 (1-2623)

QY		1	MetGluAlaThrGlyThrTrpAlaLeuLeuLeuAlaLeuLeuLeuLeuLeuThr	20
Db		41	ATGGAGGCACCGCACCTGGCGGTGCTGTGGCGTGGCGTGCTCTGCTGCTGACG	100
QY		21	LeuAlaLeuSerGlyThrArgAlaArgGlyHisLeuProGlyProThrLeuPro	40
Db		101	CTGGGCGTGTCCGGACCCAGGGCCCAGGCCACTGTCGCCCGGGCCACGCGGCTACCA	160
QY		41	LeuLeuGlyAsnLeuLeuGlnLeuArgProGlyValAlaLeuTyrsSerGlyLeuMetArgLeu	60
Db		161	CTGTGGGAACCTCTGCACCTACGGCCCCGGGGCGCTGTATTACGGGCTCATGCGGCTG	220
QY		61	SerLysLysTyrglyProValPheThrIleTyrlieuGlyProThrArgProValVal	80
Db		221	AGTAAGAAGTACGGACCGGTGTTCAACCATCTACCTGGGACCCCTGGCGGCTGTGGTGGTC	280
QY		81	LeuValGlyGlnGluAlaValArgGluAlaLeuGlyGlyGlnAlaGluGluPheSerGly	100
Db		281	CTGGTTGGCAGGAGGCTGTGCGGGAGGCCCTGGGAGGTTCAGGAGTTACGCGGC	340
QY		101	ArgGlyThrValAlaMetLeuGluGlyThrPheAspGlyHiscglyValPhePheSerAsn	120
Db		341	CGGGGACCGTAGCATGCTTGGAAAGGACATTTTATGGCCATGGGGTTTTCTCTCCAAC	400
QY		121	GlyGluArgTrpArgGlnLeuArgLysPheThrMetLeuAlaLeuArgAspLeuGlyMet	140

Db 401 GGGAGCGGTGGAGGCGAGCTGAGGAGTTTACCATGCTTGTCTCTCGGGAGCCTGGGCATG 460  
Qy 141 GlyLysArgGluGlyGluGluLeuIleGlnAlaArgCysLeuValGluThrPhe 160  
Db 461 GGGAGCGAGAGCGGAGGAGCTGATCAGCGGAGGCGCGGTGTCTGGTGGAGACATTC 520  
Qy 161 GlnGlyThrGluGlyArgProPheAspProSerLeuLeuAlaGlnAlaThrSerAsn 180  
Db 521 CAGGGGACAGAGGAGCGCCATTCGATCCCTCCCTGCTGCTGGCCAGCGCCACCTCCAAAC 580  
Qy 191 ValValCysSerLeuLeuPheGlyLeuArgPheSerTyrGluAspLysGluPheGlnAla 200  
Db 581 GTAGTCT 640  
Qy 201 ValValArgAlaAlaGlyGlyThrLeuLeuGlyValSerSerGlnGlyGlyGlnThrTyr 220  
Db 641 GTGTCTCCGGCAGCTGGTGTACCTGCTGGAGTCAGCTCCAGGGGGGTCCAGACCTAC 700  
Qy 221 GluMetPheSerTrpPheLeuArgProLeuProGlyProHisLysGlnLeuLeuHis 240  
Db 701 CAGATGTTCTCTCTGTTCTCTCGGGCCCTGCCAGGCCCCCAAGCAGGCTCTCCACCCAC 760  
Qy 241 ValSerThrLeuAlaAlaPheThrValArgGlnValGlnGlnHisGlnGlyAsnLeuAsp 260  
Db 761 GTCAGACCTTGGCTGCTCTTCACTGTCGAGTGCAGCAGCAGCAGCAGCAGCAGCAGCAG 820  
Qy 261 AlaSerGlyProAlaArgAspLeuValAspAlaPheLeuLeuLysMetAlaGlnGluGlu 280  
Db 821 GCTTCGGGCCCCGCGACCTTGTGATGCTTCTCTGTAAGATGGCAGAGGAGAA 880  
Qy 281 GlnAsnProGlyThrGluPheThrAsnLysAsnMetLeuMetThrValIleTyrLeuLeu 300  
Db 881 CAAAACCCAGGACACAGAAATTCACCAACAAGAAATCTGATGACAGTCAATTAATTGCTG 940  
Qy 301 PheAlaGlyThrMetThrValSerThrThrValGlyTyrThrLeuLeuLeuMetLys 320  
Db 941 TTTCCTGGAGCATGAGGTGACGACACACCGTGGGTATACCTCTCTCTCTGATGATAA 1000  
Qy 321 TyrProHisValGlnLysTrpValArgGluGluLeuAsnArgGluLeuGlyAlaGlyGln 340  
Db 1001 TACCCTCATGTCCAAAGTGGGTACGTGAGAGCTGAATCGGAGCTGGGGGTGGCCAG 1060  
Qy 341 AlaProSerLeuGlyAspArgThrArgLeuProTyrThrAspAlaValLeuHisGluAla 360  
Db 1061 GCACCAAGCCTAGGGACCGGTACCCGCTCCCTTACACCGACCGGTTCGTGATGAGCG 1120  
Qy 361 GlnArgLeuLeuAlaLeuValProMetGlyIleProArgThrLeuMetArgThrThrArg 380  
Db 1121 CAGCGCTGCTGGGCTGGTGGCCATGGGAATACCCGACCTCATGGGACCAACCCGC 1180  
Qy 381 PheArgGlyTyrThrLeuProGlnGlyThrGluValPheProLeuLeuGlySerIleLeu 400  
Db 1181 TTCGAGGGTACACCTTCGCCAGGCGACGAGGTCTTCCCTCTCTCTCTCTCTCTCTCT 1240  
Qy 401 HisAspProAsnIlePheLysHisProGluGluPheAsnProAspArgPheLeuAspAla 420  
Db 1241 CATGACCCCAACATCTTCAAGACACCCAGAGAGTTCAACCCAGACCGTTTCTTGGATGCA 1300  
Qy 421 AspGlyArgPheArgLysHisGluAlaPheLeuProPheSerLeuGlyLysArgValCys 440  
Db 1301 GATGACGGTTCAGNAAGCATGAGCGTCTCTGCTCTCTCTCTCTCTCTCTCTCTCTCT 1360  
Qy 441 LeuGlyGluGlyLeuAlaLysAlaGluLeuPheLeuPhePheThrThrIleLeuGlnAla 460  
Db 1361 CTTGGAGAGGCGCTGGAAAAAGCGAGCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1420  
Qy 461 PheSerLeuGluSerProCysProAspThrLeuSerLysLeuSerLysProThrValSerGly 480  
Db 1421 TTCTCTCTGGAGAGCCCGTCCCGCGGACACCTCTGAGCCTCAAGCCACCGCTCAGTGGC 1480  
Qy 481 LeuPheAsnIleProProAlaPheGlnLeuGlnValArgProThrAspLeuHisSerThr 500  
Db 1481 CTTTTCACATTCCTCCCGAGCCTTCCAGCTGCAAGTCCGCTGCCACTGACCTTCACTCCACC 1540

Qy 501 ThrGlnThrArg 504  
Db 1541 ACGCAGACCAGA 1552  
RESULT 15  
US-10-144-993-63  
; Sequence 63, Application US/10144993  
; Publication No. US20040038336A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: DeForge, Laura  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Sherwood, Steven  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3330R1C261  
; CURRENT APPLICATION NUMBER: US/10/144,993  
; CURRENT FILING DATE: 2002-05-13  
; Prior Application removed - See File Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 550  
; SEQ ID NO 63  
; LENGTH: 2623  
; TYPE: DNA  
; ORGANISM: Homo Sapien  
US-10-144-993-63  
Alignment Scores: Length: 5,88e-314 2623  
Pred. No.: 2615.00 504  
Score: 2615.00  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 100.00%  
DB: 13 Gaps: 0  
US-10-669-693-2 (1-504) x US-10-144-993-63 (1-2623)  
Qy 1 MetGluAlaThrGlyThrTrpAlaLeuLeuLeuAlaLeuLeuLeuLeuLeuLeuThr 20  
Db 41 ATGAGGGGACCGGACCTGGGCGCTGCTGCGCTGGCGCTGCTCTCTCTCTCTCTCTCAGC 100  
Qy 21 LeuAlaLeuSerGlyThrArgAlaArgGlyHisLeuProGlyProGlyProThrProLeuPro 40  
Db 101 CTGGCGCTGTCCGGGACCGAGGCGCCACCTGCCCCCGGCGCCACCGCGCTACCA 160  
Qy 41 LeuLeuGlyAsnLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 60  
Db 161 CTCTGGGAAACCTCTCTGCGAGTACGCGCCCGGCGCGCTGATTTCAGGGCTCATCGCGCTG 220  
Qy 61 SerLysLysTyrGlyProValPheThrIleTyrLeuGlyProTrpArgProValValVal 80  
Db 221 AGTAAGAAGTACGAGCCCGGTTTCAACCATCTACTGGGACCTTGGCGCGCTGTGGTGGTC 280  
Qy 81 LeuValGlyGlnGluAlaValArgGluAlaLeuGlyGlyGlnAlaGluGluPheSerGly 100  
Db 281 CTGTTTGGGAGGAGGCTGTGCGGAGGCGCTTGGAGGTGAGGTGAGGTGAGGTGAGGTGAGGT 340  
Qy 101 ArgGlyThrValAlaMetLeuGluGlyThrPheAspGlyHisGlyValPhePheSerAsn 120  
Db 341 CGGGGAAACCGTAGCATGCTGGAGAGGACCTTTGATGGCCATGGGGTTTCTTCTCTCCAAC 400

```
QY 121 GlyGluArgTTPArgGlnLeuArgLysPheThrMetLeuAlaLeuArgAspLeuGlyMet 140
Db |||||
QY 401 GGGGAGCGGTGGAGCGAGCTGAGGAAGATTTACATGCTTGGCTTCTCGGGACCTGGGCGATG 460
Db |||||
QY 141 GlyLysArgGluGlyGluLeuLeuGlnAlaGluAlaArgCysLeuValGluThrPhe 160
Db |||||
QY 461 GGGAAAGCGAAGGCGAGGAGCTGATCCAGGCGGAGGCCCGGTGTCGGTGGAGACATTC 520
Db |||||
QY 161 GlnGlyThrGluGlyArgProPheAspProSerLeuLeuAlaGlnAlaThrSerAsn 180
Db |||||
QY 521 CAGGGGACAGAAAGGACGCCCATTCGATCCCTCCCTGCTGCTGGCCACAGGCCACCTCCCAAC 580
Db |||||
QY 181 ValValCysSerLeuLeuPheGlyLeuArgPheSerTyrGluAspLysGluPheGlnAla 200
Db |||||
QY 581 GTAGTCTGCTCCCTCTCTTTGGCCCTCGCTTCTCTATGAGGATAAGGAGTTCAGGCC 640
Db |||||
QY 201 ValValArgAlaAlaGlyGlyThrLeuLeuGlyValSerSerGlnGlyGlyGlnThrTyr 220
Db |||||
QY 641 GTGGTCCGGGCGAGCTGGTGTACCTGCTGGGAGTCAGCTCCAGGGGGGTCCAGACCTAC 700
Db |||||
QY 221 GluMetPheSerTrpPheLeuArgProLeuProGlyProHisLysGlnLeuLeuHisHis 240
Db |||||
QY 701 GAGATGTTCTCTCTGCTTCTCGGGCCCTGCGAGCCGCCCAAGCAGCTCTCTCCACCAC 760
Db |||||
QY 241 ValSerThrLeuAlaAlaPheThrValArgGlnValGlnGlnHisGlnGlyAsnLeuAsp 260
Db |||||
QY 761 GTCAGCACCTTGGCTGCTTCACAGTCCGGCGAGGTGCAGCAGCACCCAGGGGAACCTGGAT 820
Db |||||
QY 261 AlaSerGlyProAlaArgAspLeuValAspAlaPheLeuLeuLysMetAlaGlnGluGlu 280
Db |||||
QY 821 GCTTCGGGGCCCCGACGCTGACTGTGATGCCCTTCTGCTGAAGATGGCACAGAGGAA 880
Db |||||
QY 281 GlnAsnProGlyThrGluPheThrAsnLysAsnMetLeuMetThrValIleTyrLeuLeu 300
Db |||||
QY 881 CAAAAACCAGGCACAGAATTCCAAACAAGAACATGCTGATGACAGTCATTTATTGCTG 940
Db |||||
QY 301 PheAlaGlyThrMetThrValSerThrThrValGlyTyrThrLeuLeuLeuMetLys 320
Db |||||
QY 941 TTTGCTGGGACGATGACGGTCAGCACCGTCGGCTATACCCCTCTGCTCTGCTGATGAAA 1000
Db |||||
QY 321 TyrProHisValGlnLysTrpValArgGluLeuLeuAsnArgGluLeuGlyValaGlyGln 340
Db |||||
QY 1001 TACCCCTCATGTCCAAAAGTGGGTACGTGAGGAGCTGAATCGGGAGCTGGGGGCTGGCCAG 1060
Db |||||
QY 341 AlaProSerLeuGlyAspArgThrArgLeuProTyrThrAspAlaValLeuHisGluAla 360
Db |||||
QY 1061 GCACCAAGCCCTAGGGGACCGTACCCGCTCTCCCTTACACCGACGCGTTCGTCATGAGCG 1120
Db |||||
QY 361 GlnArgLeuLeuAlaLeuValProMetGlyLeProArgThrLeuMetArgThrThrArg 380
Db |||||
QY 1121 CAGCGGCTGCTGGGCTGGTGGCCATGGGNAATACCCCGCACCCCTCATGGCGACCCCCGC 1180
Db |||||
QY 381 PheArgGlyTyrThrLeuProGlnGlyThrGluValPheProLeuLeuGlySerIleLeu 400
Db |||||
QY 1181 TTCGAGGGTACACCTCGCCCGCCAGGGCAGCGAGTCTTCCCTCTCTCTGCTTGCATCCTG 1240
Db |||||
QY 401 HisAspProAsnIlePheLysHisProGluGluPheAsnProAspArgPheLeuAspAla 420
Db |||||
QY 1241 CATGACCCCAACATCTCAAGACACCCAGAAAGAGTTCAACCCAGACCGTTTCTCTGGATGCA 1300
Db |||||
QY 421 AspGlyArgPheArgLysHisGluAlaPheLeuProPheSerLeuGlyLysArgValCys 440
Db |||||
QY 1301 GATGGACGGTTTCAGGAAGCATGAGGGGTTCCTGCCCTTCTCTTAGGGAAGCGGTCTGC 1360
Db |||||
QY 441 LeuGlyGluGlyLeuAlaLysAlaGluLeuPhePheThrThrIleLeuGlnAla 460
Db |||||
QY 1361 CTTGGAGAGGGCCCTGGCAAAAGCGGAGCTCTTCTCTCTTTCACCACTCTTACAGGCC 1420
Db |||||
QY 461 PheSerLeuGluSerProCysProAspThrLeuSerLeuLysProThrValSerGly 480
Db |||||
QY 1421 TTCTCTCTGGAGAGCCGCTGCCCGCGGACACCTTGAGCCTCAAGCCCAACCGTCAGTGGC 1480
Db |||||
```

```
QY 481 LeuPheAsnIleProProAlaPheGlnLeuValArgProThrAspLeuHisSerThr 500
Db |||||
QY 1481 CTTTTTCAACATTTCCCCCAGCCTTCCAGCTTCCAGTCCGTCCTCCACTGACCTTCACTCCACC 1540
Db |||||
QY 501 ThrGlnThrArg 504
Db 1541 ACGCAGACCAGA 1552
```

Search completed: September 18, 2004, 06:17:48  
Job time : 777 secs

***This Page Blank (uspto)***



GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 17, 2004, 18:58:32 ; Search time 635 Seconds  
(without alignments)  
3371.798 Million cell updates/sec

Title: US-10-669-693-2  
Perfect score: 2615  
Sequence: 1 MEATGTWALLALALLLLT.....PPAFQLQVRPTDLHSTTQTR 504

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2.1/USPTO.spool/US10669693/runat.15092004.101309.5839/app.query.fasta.1.647  
-DB=N Geneseq 29Jan04 -QFMR=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10669693 @CGN 1.1 470 @runat.15092004.101309.5839 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N\_Geneseq 29Jan04:\*

1: Geneseqn1980s:\*

2: Geneseqn1990s:\*

3: Geneseqn2000s:\*

4: Geneseqn2001as:\*

5: Geneseqn2001bs:\*

6: Geneseqn2002s:\*

7: Geneseqn2003as:\*

8: Geneseqn2003bs:\*

9: Geneseqn2003cs:\*

10: Geneseqn2004as:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	2615	100.0	1515	6	ABT05472
2	2615	100.0	1527	9	ACA61907
3	2615	100.0	1927	7	ADA60552
4	2615	100.0	2085	7	ABX71208
5	2615	100.0	2604	6	ABQ72550
6	2615	100.0	2620	6	AA224671
7	2615	100.0	2623	4	AA221275
8	2615	100.0	2623	7	ACD23884

9	2615	100.0	2623	7	ACA67025
10	2615	100.0	2623	7	ACA03634
11	2615	100.0	2623	7	ABX89172
12	2615	100.0	2623	7	ACD41826
13	2615	100.0	2623	7	ACA04055
14	2615	100.0	2623	8	ADA45582
15	2615	100.0	2623	8	ADA76013
16	2615	100.0	2623	8	ADA18663
17	2615	100.0	2623	8	ADA61286
18	2615	100.0	2623	8	ADB19071
19	2615	100.0	2623	8	ADB27612
20	2615	100.0	2623	8	ADA86091
21	2615	100.0	2623	8	ADB15655
22	2615	100.0	2623	8	ADA47441
23	2615	100.0	2623	8	ADA67236
24	2615	100.0	2623	8	ADB30243
25	2615	100.0	2623	8	ADA85539
26	2615	100.0	2623	8	ADA87551
27	2615	100.0	2623	8	ADA79055
28	2615	100.0	2623	8	ADA87194
29	2615	100.0	2623	8	ADB16396
30	2615	100.0	2623	8	ADA91488
31	2615	100.0	2623	8	ADB14551
32	2615	100.0	2623	8	ADB18512
33	2615	100.0	2623	8	ADA93727
34	2615	100.0	2623	8	ADB19623
35	2615	100.0	2623	8	ACD98455
36	2615	100.0	2623	8	ADA74189
37	2615	100.0	2623	8	ADA84435
38	2615	100.0	2623	8	ADA84987
39	2615	100.0	2623	8	ADA81946
40	2615	100.0	2623	8	ADA74909
41	2615	100.0	2623	8	ADA84987
42	2615	100.0	2623	8	ADA84435
43	2615	100.0	2623	8	ADB29691
44	2615	100.0	2623	8	ADA80219
45	2615	100.0	2623	8	ADA75461

#### ALIGNMENTS

RESULT 1

ABT05472

ID ABT05472 standard; DNA; 1515 BP.

XX

AC ABT05472;

XX

DT 11-OCT-2002 (first entry)

XX

DE DNA of NOVX 16 SEQ ID No 39.

XX

Cytostatic; antidiabetic; anorectic; metabolic; neurotropic; antilipemic; neuroprotective; antiparkinsonian; anticonvulsant; cerebroprotective; tranquiliser; neuroleptic; antidiabetic; antitumor; antineoplastic; anti-HIV; antiallergic; antirheumatic; antithrombotic; antidiabetic; metabolic disorder; obesity; infectious disease; Alzheimer's disease; anorexia; neurodegenerative disorder; Parkinson's disease; obesity; immune disorder; haematopoietic disorder; dyslipidaemia; chronic disease; metabolic syndrome X; wasting disorder; cancer; neurological disorder; epilepsy; stroke; mental disorder; schizophrenia disorders; goiter; vesicular transport; cystic fibrosis; gastrointestinal disorder; diabetes mellitus; ulcerative colitis; AIDS; allergic reaction; multiple sclerosis; rheumatoid arthritis; transgenic animal; gene therapy; gene; ds.

XX Unidentified.

OS

XX WO200246409-A2.

XX

XX 13-JUN-2002.

XX

PF 06-DEC-2001; 2001WO-US046586.

XX





181 ValValCysSerLeuLeuPheGlyLeuArgPheSerTyrGluAspLysGluPheGlnAla 200  
541 GTAGTCTGCTCCCTCTCTTTGGCTCGCTTCTCTATGAGGATAAGGAGTTCCAGGCC 600  
201 ValValArgAlaAlaGlyThrLeuLeuGlyValSerSerGlnGlyGlnThrTyr 220  
601 GTGGTCCGGGCGAGTGTGTGTACCTCTGTGGAGTCAGCTCCAGGGGGGTCAGACCTAC 660  
221 GluMetPheSerTrpLeuArgProLeuProGlyProHisLysGlnLeuLeuHisHis 240  
661 GAGATGTCTCTCTGCTTCTCTGGGCGCTGTGACAGGCCCCCAGACGAGCTCTCCACCAC 720  
241 ValSerThrLeuAlaAlaPheThrValArgGlnValGlnGlnHisGlnGlyAsnLeuAsp 260  
721 GTGAGCACCTTGGCTGGCTTTCACAGTCCGCGAGGTGCAGCAGCACAGGGGAACTGGAT 780  
261 AlaSerGlyProAlaArgAspLeuValAspAlaPheLeuLeuLeuMetAlaGlnGluGlu 280  
781 GCTTTCGGGCGCCGACGTCGACCTTGTGATGCCCTTCTGCTGAAGATGGCAGAGGAA 840  
281 GlnAsnProGlyThrGluPheThrAsnLysAsnMetLeuMetThrValIleTyrLeuLeu 300  
841 CAAGAACCAGGCACAGAAATTCACCAACAGACATGCTGATGACAGTCATTATTGCTG 900  
301 PheAlaGlyThrMetThrValSerThrThrValGlyTyrThrLeuLeuLeuMetLys 320  
901 TTTGCTGGGACGATGACGTCAGCAGCACCGCTCGGCTATACCTCTCTCTCTGATGAAA 960  
321 TyrProHisValGlnLysTrpValArgGluGluLeuAsnArgGluLeuGlyValAlaGlyGln 340  
961 TACCCTCATGTCCAAAGTGGGTAGCTGAGGAGCTGAATCGGGAGCTGGGGGCTGGCCAG 1020  
341 AlaProSerLeuGlyAspArgThrArgLeuProTyrThrAspAlaValLeuHisGluAla 360  
1021 GACCAAGCCTAGGGACCGTACCGCTCCCTTACACCGACGCGTCTTGCAATGAGGGG 1080  
361 GlnArgLeuLeuAlaLeuValProMetGlyIleProArgThrLeuMetArgThrThrArg 380  
1081 CAGCGCTGCTGGCGTGTGGTCCATGGGAATACCCCGCACCTCATGCGGACCAACCCGC 1140  
381 PheArgGlyTyrThrLeuProGlnGlyThrGluValPheProLeuLeuGlySerIleLeu 400  
1141 TTCCGAGGGTACACCTGCGCCAGGCGACGGAGGTCTTCCCTCTCTTGGCTCCATCTTG 1200  
401 HisAspProAsnIlePheLysHisProGluGluPheAsnProAspArgPheLeuAspAla 420  
1201 CATGACCCCAACATCTTCAAGCACCCAGAGAGTTCAACCCAGACCGTTTCTGTGGATGCA 1260  
421 AspGlyArgPheArgLysHisGluAlaPheLeuProPheSerLeuGlyLysArgValCys 440  
1261 GATGGACGGTTACGGAAGCATGAGGGGTCTCTGCTCTCTCTTACGGAAGCGTGTCTGC 1320  
441 LeuGlyGluGlyLeuAlaLysAlaGluLeuPhePheThrThrIleLeuGlnAla 460  
1321 CTTGGAGAGGGCTGGGAAAAGCGGAGCTTCTCTTCTTCAACCATCTCTACAGGCC 1380  
461 PheSerLeuGluSerProCysProProAspThrLeuSerLeuLysProThrValSerGly 480  
1381 TTCTCTCTGGAGAGCCGCTGCGCCCGGACACCTGTAGCCTCAAGCCACCGTCAGTGGC 1440  
481 LeuPheAsnIleProProAlaPheGlnLeuGlnValArgProThrAspLeuHisSerThr 500  
1441 CTTTCAACATTCGCCAGGCTTCCAGTTCGAGTCCGTCGATCCACTGACCTTCACTCCACC 1500  
501 ThrGlnThrArg 504  
1501 ACGCAGACCAGA 1512

RESULT 3  
AAD60552  
ID AAD60552 standard; cDNA; 1927 BP.  
XX

AC AAD60552;  
XX 18-DEC-2003 (first entry)  
XX Human cytochrome p450, 33303 cDNA.  
DE  
XX Human; cytochrome p450; dehydrogenase/reductase; osteomalacia; psoriasis;  
KW lipoxigenase; hydratase; proliferative disorder; haematopoietic disorder;  
KW differentiative disorder; carcinoma; sarcoma; leukaemia; immune disorder;  
KW anorexia nervosa; neural degeneration; muscular disorder; lipid disorder;  
KW multiple sclerosis; encephalomyelitis; metabolic disorder; osteoporosis;  
KW myasthenia gravis; bone metabolism; immunomodulator; obesity; anorectic;  
KW eating disorder; osteodystrophy; arthritis; diabetes; diabetes; rickets;  
XX milk fever; diabetes mellitus; gene; ss.  
OS Homo sapiens.  
XX  
XX Key Location/Qualifiers  
CDS 1..1515  
FT /\*tag= a  
FT /product= "Human 33303"  
FT /note= "CDS is specifically claimed as SEQ ID NO: 6 in  
FT claim 1 of the specification"  
FT sig\_peptide 1..87  
FT /\*tag= b  
FT mat\_peptide 88..1512  
FT /\*tag= c  
FT /product= "Mature human 33303"  
XX  
XX US2003092658-A1.  
XX  
XX 15-MAY-2003.  
XX  
XX 20-JUN-2002; 2002US-00175696.  
XX  
XX 02-FEB-2001; 2001US-0266140P.  
XX 04-FEB-2002; 2002US-00067669.  
XX  
XX (MEYE/) MEYERS R E.  
XX (GLUC/) GLUCKSMANN M A.  
XX (RUDO/) RUDOLPH-OWEN L A.  
XX  
XX Meyers RE, Glucksmann MA, Rudolph-Owen LA;  
XX  
XX WPI: 2003-765490/72.  
XX P-PSDB: AAE39880.  
XX  
XX New cytochrome P450 family members, dehydrogenase/reductase protein,  
XX lipoxigenase family member and human hydratase useful for treating  
XX cellular proliferative and/or differentiative disorders.  
XX  
XX Claim 1; Page 170-171; Opp; English.  
XX  
XX The invention relates to isolated cytochrome p450 family members such as  
XX 33312, 33302 and 32579, dehydrogenase/reductase protein such as 21509,  
XX 33770, lipoxigenase family member such as 46638 and human hydratase such  
XX as 50090. The invention is useful as diagnostic targets and agents for  
XX controlling one or more of cellular proliferative and/or differentiative  
XX disorders such as carcinoma, sarcoma or haematopoietic disorders such as  
XX leukaemia, immune disorders such as diabetes mellitus, arthritis,  
XX multiple sclerosis, encephalomyelitis, myasthenia gravis, psoriasis, etc.  
XX or metabolic disorders such as obesity, anorexia nervosa, cachexia, lipid  
XX disorders and diabetes. The invention is useful for controlling disorders  
XX associated with abnormal fatty acid biosynthesis and neural degeneration.  
XX The invention is useful for controlling disorders associated with defects  
XX in fatty acid oxidation or proliferation or muscular disorders and for  
XX treating disorders associated with bone metabolism such as osteoporosis,  
XX osteodystrophy, osteomalacia, rickets, osteitis fibrosa cystica, milk  
XX fever. The present sequence is human cytochrome p450 cDNA  
XX  
XX Sequence 1927 BP; 383 A; 625 C; 550 G; 369 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	3.59e-231	Length:	1927
Score:	2615.00	Matches:	504
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0

US-10-669-693-2 (1-504) x AAD60552 (1-1927)

[illegible]

QY	321	TyrProHisValGlnLysTyrPValArgGluGluLeuAsnArgGluLeuGlyValAlaGlyCln	340
Db	961	TACCCCTCATGTCCTCAAAAGTGCGTACGTGAGGAGCTGAATCGGAGAGCTGGGGGCTGGCGAG	1020
QY	341	AlaProSerLeuGlyAspArgThrArgLeuProTyrThrAspAlaValLeuHisGluAla	360
Db	1021	GCACCAAGCCTTAGGGGACCGTACCGCGCTCCCTTACACCGACGCGGTTCCTGCATGAGCGC	1080
QY	361	GlnArgLeuLeuAlaLeuValProMetGlyIleProArgThrLeuMetArgThrThrArg	380
Db	1081	CAGCGGCTGCTGGCGCTGGTGCCCATGGGAATACCCGACCCCTCATGCGGACCAACCCGC	1140
QY	381	PheArgGlyTyrThrLeuProGlnGlyThrGluValPheProLeuLeuGlySerIleLeu	400
Db	1141	TTCCGAGGGTACACCCCTGCCACGGGACGGAGGTCTTCCCTCTCTTGGCTTCATCCTG	1200
QY	401	HisAspProAsnIlePheLysHisProGluGluPheAsnProAspArgPheLeuAspAla	420
Db	1201	CATGACCCCAACATCTTCAAGCACCCAGAAAGATTCAACCCAGACCGTTTCTCTGGATGCA	1260
QY	421	AspGlyVargPheArgLysHisGluAlaPheLeuProPheSerLeuGlyLysArgValCys	440
Db	1261	GATGGACGGTTTCAGGAAGCATGAGGCGTCTCTCCCTCTCTTCTTAGGAAGCGTGTCTGC	1320
QY	441	LeuGlyGluGlyLeuAlaLysAlaGluLeuPhePhePheThrThrIleLeuGlnAla	460
Db	1321	CTTGGAGAGGCGCTGGCAAAAGCGGAGCTTCTCTCTTTTACCAACCATTCTTACAAGCC	1380
QY	461	PheSerLeuGlnSerProCysProProAspThrLeuSerLeuLysProThrValSerGly	480
Db	1381	TTCTCCCTGGAGAGCCGTGTCGCCCGCGGACACCTGAGCCTCAAGCCCAACCGTCAGTGCC	1440
QY	481	LeuPheAsnIleProProAlaPheGlnLeuGlnValArgProThrAspLeuHisSerThr	500
Db	1441	CTTTTCAACATCTCCCAAGCCTTCAGAGTCGAAGTCCGTCACCTGACCTTCACTCCACC	1500
QY	501	ThrGlnThrArg	504
Db	1501	ACGCAGACCAGA	1512
RESULT 4			
ABX71208			
ID	ABX71208 standard; cDNA; 2085 BP.		
XX	AC ABX71208;		
XX	AC ABX71208;		
DT	05-MAR-2003 (first entry)		
XX	Novel human cDNA sequence #433.		
XX	Human; Gene; ss; nervous system disorder; peripheral neuropathy;		
KW	Huntington's disease; amyotrophic lateral sclerosis; haemophilia;		
KW	neurodegenerative disease; Parkinson's disease; Alzheimer's disease;		
KW	autoimmune disease; systemic lupus erythematosus; rheumatoid arthritis;		
KW	insulin-dependent diabetes mellitus; anaemia; thrombocytopaenia; wound;		
KW	ulcer; burn; bone disorder; osteoporosis; osteoarthritis; stroke;		
KW	fibrosis; reperfusion injury; infection; allergic rhinitis; asthma;		
KW	coagulation disorder; cancer; tumour; inflammatory disease; septic shock;		
KW	Crohn's disease; anaphylaxis; proliferation; chemotactic;		
KW	differentiation; stem cell growth factor; haematopoiesis; chemokinetic;		
KW	haemostatic; antiinflammatory; expressed sequence tag, EST.		
OS	Homo sapiens.		
XX	OS		
PN	WO200281731-A2.		
XX	17-OCT-2002.		
XX	29-JAN-2002; 2002WO-US001222.		
PF	XX		
XX	30-JAN-2001; 2001US-00774528.		
FR	(HYSE-) HYSEO INC.		
PA	XX		







1480 CTTTTCACATTCCTCCAGCCTTCAGCTGCAAGTCCGTCCTCCACTGACCTTCACCTCCACC 1539

Db 400 GGGGAGCGGTGGAGCGAGCTGAGGAGTTTACCATGCTTGTCTGCGGAGCCTGGGCAATG 459  
 Qy 141 GLVLSARGGLUGLUGLUGLULEGLNAlaGLUAlaAraCysLeuValGluThrPhe 160  
 Db 460 GGGAGCGAGAGGAGGAGAGCTGATCCAGCGGAGGCGCGGTGCTGTGTGGAGACATTC 519  
 Qy 161 GlnGlyThrGluGlyArgProPheAspProSerLeuLeuLeuAlaGlnAlaThrSerAsn 180  
 Db 520 CAGGGGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 579  
 Qy 191 ValValCysSerLeuLeuPheGlyLeuArgPheSerTyrGluAspLysGluPheGlnAla 200  
 Db 580 GTAGTCT 639  
 Qy 201 ValValArgAlaAlaGlyGlyThrLeuLeuGlyValSerSerGlnGlyGlnThrTyr 220  
 Db 640 GTGGTCCGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 699  
 Qy 221 GluMetPheSerTrpPheLeuArgProLeuProGlyProHisLysGlnLeuLeuHis 240  
 Db 700 GAGATGTTCTCTGTGTTCTGTGGGCGCTGCCAGGCGCCCAACAGCAGCTCTCCACCA 759  
 Qy 241 ValSerThrLeuAlaAlaPheThrValArgGlnValGlnGlnHisGlnGlyAsnLeuAsp 260  
 Db 760 GTCAGCACCTTGGCTGCTTCACAGTCCGCGAGGTCAGCAGCAGCAGGAGACCTGGAT 819  
 Qy 261 AlaSerGlyProAlaArgAspLeuValAspAlaPheLeuLeuLysMetAlaGlnGlu 280  
 Db 820 GTTTCGGGCGCGCAGCTGACCTTGTGATGCTTCTCTCTCTCTCTCTCTCTCTCTCT 879  
 Qy 281 GlnAsnProGlyThrGluPheThrAsnLysAsnMetLeuMetThrValIleTyrLeuLeu 300  
 Db 880 CAAACCCAGGACAGAAATTCACCAACAAAGACATGCTGATGACAGTCAATTTATTGCTG 939  
 Qy 301 PheAlaGlyThrMetThrValSerThrThrValGlyTyrThrLeuLeuLeuMetLys 320  
 Db 940 TTTGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 999  
 Qy 321 TyrProHisValGlnLysTrpValArgGluGluLeuAsnArgGluLeuGlyAlaGlyGln 340  
 Db 1000 TACCCTCATGTCACAAAGTGGTACGTGAGGAGCTGAATCGGAGCTGGGGGCTGGCCAG 1059  
 Qy 341 AlaProSerLeuGlyAspArgThrArgLeuProTyrThrAspAlaValLeuHisGluAla 360  
 Db 1060 GCACCAAGGCTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1119  
 Qy 361 GlnArgLeuLeuAlaLeuValProMetGlyLeuProArgThrLeuMetArgThrThrArg 380  
 Db 1120 CAGCGGCTGTGGCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1179  
 Qy 381 PheArgGlyTyrThrLeuProGlnGlyThrGluValPheProLeuLeuGlySerIleLeu 400  
 Db 1180 TTCCGAGGTTACACCTGCCCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1239  
 Qy 401 HisAspProAsnIlePheLysHisProGluGluPheAsnProAspArgPheLeuAspAla 420  
 Db 1240 CATGACCCCAACATCTTCAAGACCCCAAGAGAGTTCACCCAGACCGTTCCTTGGATGCA 1299  
 Qy 421 AspGlyArgPheArgLysHisGluAlaPheLeuProPheSerLeuGlyLysArgValCys 440  
 Db 1300 GATGGAGGTTTACAGAGCATGAGCGGTTCTGCGCTTCTCTTACGAGGAGGAGGAGG 1359  
 Qy 441 LeuGlyGluGlyLeuAlaLysAlaGluLeuPhePheThrThrIleLeuGlnAla 460  
 Db 1360 CTTGGAGGCGCTGGCAAAAGCGAGCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1419  
 Qy 461 PheSerLeuGluSerProCysProAspThrLeuSerLeuLysProThrValSerGly 480  
 Db 1420 TTCTCCCTGGAGAGCGCGGCGCGGAGACCTGAGCCCTCAAGCCACCGCTCAGTGGC 1479  
 Qy 481 LeuPheAsnIleProProAlaPheGlnLeuGlnValArgProThrAspLeuHisSerThr 500

Db 1480 CTTTTCACATTCCTCCAGCCTTCAGCTGCAAGTCCGTCCTCCACTGACCTTCACCTCCACC 1539  
 Qy 501 ThrGlnThrArg 504  
 Db 1540 ACGCAGACCAGA 1551  
 RESULT 6  
 ID AAD24671 standard; cDNA; 2620 BP.  
 AC AAD24671;  
 XX 12-MAR-2002 (first entry)  
 DT Human drug metabolising enzyme (DME)-6 cDNA.  
 DE Human; drug metabolising enzyme; gene therapy; autoimmune disorder;  
 KW inflammatory disorder; acquired immune deficiency syndrome; infection;  
 KW AIDS; adult respiratory distress syndrome; cell proliferative disorder;  
 KW allergy; anaemia; conjunctivitis; actinic keratosis; arteriosclerosis;  
 KW cancer; endocrine disorder; hypothalamus disorder; pituitary disorder;  
 KW gastrointestinal disorder; metabolic disorder; developmental disorder;  
 KW liver disorder; iritis; cystic fibrosis; Addison's disease; retinitis;  
 KW goitre; diabetes; dysphagia; gastric carcinoma; anorexia; eye disorder;  
 KW DME-6; ss.  
 OS Homo sapiens.  
 XX Key Location/Qualifiers  
 FH CDS 41..1555  
 FT /tag= a  
 FT /product= "Human DME-6 protein"  
 FT sig\_peptide 41..124  
 FT /tag= b  
 FT mat\_peptide 125..1552  
 FT /tag= c  
 FT /product= "Human mature DME-6 protein"  
 XX WO200179468-A2.  
 PD 25-OCT-2001.  
 XX 12-APR-2001; 2001WO-US011869.  
 XX 13-APR-2000; 2000US-0197590P.  
 PR 19-APR-2000; 2000US-0198403P.  
 PR 28-APR-2000; 2000US-0200185P.  
 PR 05-MAY-2000; 2000US-0202234P.  
 PR 11-MAY-2000; 2000US-0203509P.  
 XX (INCY-) INCYTE GENOMICS INC.  
 PA Policky JL, Hafalia A, Burford N, Ring HZ, Lal P, Tribouley CM;  
 PI Yao MG, Yue H, Tang YT, Patterson C, Das D, Sanjanwala MS;  
 PI Gandhi AR, Reddy R, Khan FA, Baughn MR, Ramkumar J, Griffin JA;  
 PI Au-Young J;  
 XX WPI; 2002-066363/09.  
 DR P-PSDB; AAE15439.  
 XX Novel isolated human drug metabolizing enzymes referred as DME 1-10,  
 PT useful for diagnosing, treating, or preventing disorders associated with  
 PT aberrant expression of DME such as allergy, anemia, asthma, infertility.  
 XX Claim 5; Page 139-140; 143pp; English.  
 XX The invention relates to human drug metabolising enzymes referred as DME  
 CC and nucleic acid molecules encoding such enzymes. Polynucleotides of the  
 CC invention are useful for assessing toxicity of test compounds and in gene  
 CC therapy. Sequences of the invention are useful in the diagnosis,  
 CC prevention and treatment of autoimmune/inflammatory disorders such as  
 CC acquired immune deficiency syndrome (AIDS), adult respiratory distress  
 CC syndrome, allergies, anaemia, atherosclerosis, asthma, autoimmune

XX

polypeptides, to link bioactive molecules to cells expressing PRO  
polypeptides, to modulate biological activities of cells expressing PRO  
polypeptides, and to detect the presence of mammalian lung, colon,  
breast, prostate, rectal, cervical or liver tumours by comparing PRO  
polypeptide expression in a cell sample to that in a control sample. Some  
of the 275 sequences are also useful to stimulate the release of tumour  
necrosis factor-alpha (TNF-alpha) from human blood, the proliferation or  
differentiation of chondrocytes, the proliferation or gene expression in  
pericyte cells, the release of proteoglycans from cartilage, the  
proliferation of inner ear utricular supporting cells or of T-  
lymphocytes, the release of a cytokine from peripheral blood monocytes  
(PBMCs), or the proliferation of endothelial cells. Some of the PRO  
polypeptides may modulate glucose or free fatty acid uptake by skeletal  
muscle cells or by adipocytes; or inhibit binding of A-peptide to factor  
VIIA. The PRO polypeptides can be used in assays to identify molecules  
involved in binding interactions. The polynucleotides encoding PRO  
polypeptides can be used to generate probes, antisense RNA/DNA,  
transgenic or knock out animals and can be used in gene therapy  
XX Sequence 2623 BP; 552 A; 865 C; 704 G; 502 T; 0 U; 0 Other;  
SQ  
Alignment Scores:  
Pred. No.: 5,37e-231 Length: 2623  
Score: 2615.00 Matches: 504  
Percent Similarity: 100.00% Conservatives: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0  
US-10-669-693-2 (1-504) x AAS21275 (1-2623)  
QY 1 MetGluAlaThrGlyThrAlaLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuThr 20  
Db 41 ATGGAGGCGACCGGCACCTGGCGCTGCTGCTGGCGCTGCTGCTGCTGCTGAGC 100  
QY 21 LeuAlaLeuSerGlyThrArgAlaArgGlyHisLeuProProGlyProThrProLeuPro 40  
Db 101 CTGGCGCTGTCCGGGACCGAGGCGCCGAGCCACCTGCCCGCGCGCCGCGCGCTACCA 160  
QY 41 LeuLeuGlyValAsnLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 60  
Db 161 CTGCTGGGAAACCTCTCTGAGCTACGCGCGCGCGCGCGCTGTATTCAGGGCTCATGGCGCTG 220  
QY 61 SerLysLysThrGlyProValPheThrIleTyrLeuGlyProThrProValValVal 80  
Db 221 AGTAAGAAGTACGAGCGCGGTGTTACCACTACTACCTGGGACCCCTGGCGCGCTGTGGTGC 280  
QY 81 LeuValGlyGlnGlnAlaValArgGluAlaLeuGlyGlyGlnAlaGluGluPheSerGly 100  
Db 281 CTGGTTGGGCGAGGAGCTGTGGCGGAGGCGCTGGGAGGCTCAGGCTGAGGAGTTCAGCGGC 340  
QY 101 AtgGlyThrValAlaMetLeuGluGlyThrPheAspGlyHisGlyValPhePheSerAsn 120  
Db 341 CGGGGAACCGTAGCATGCTGGGAAGGACCTTTTGTATGGCCATGGGTTTTTCTTCTCCAAC 400  
QY 121 GlyGluArgTrpArgGlnLeuArgLysPheThrMetLeuAlaLeuArgAspLeuGlyMet 140  
Db 401 GGGGAGCGGTGGAGGAGCTGAGGAAAGTTTACCATGCTTCTCTGCGGGACCTGGGGCATG 460  
QY 141 GlyIleArgGluGlyGluGluLeuLeuGlnAlaGluAlaArgCysLeuValGluThrPhe 160  
Db 461 GGGGAACGAGAGCGAGGAGCTGATCCAGCGGAGGCGCGGTGTCTGTGGTGGAGACATTC 520  
QY 161 GlnGlyThrGluGlyArgProPheAspProSerLeuLeuLeuAlaGlnAlaThrSerAsn 180  
Db 521 CAGGGGACAGAGGAGCGCCATTCGATCCCTCTGCTGCTGGCCAGGCGCACCTCCAAC 580  
QY 181 ValValCysSerLeuLeuPheGlyLeuArgPheSerTyrGluAspLysGluPheGlnAla 200  
Db 581 GTAGTCTGCTCCCTCTCTTTGGCCCTCCCTTCTCTATGAGGATAGGAGTTCAGGCC 640  
QY 201 ValValArgAlaAlaGlyGlyThrLeuLeuGlyValSerSerGlnGlyGlyGlnThrTyr 220

DE Human cDNA sequence encoding for PRO1906 polypeptide.  
XX Human secretory and transmembrane; PRO; mammalian; cancer; lung; breast;  
KW prostate; cervical; tumour necrosis factor-alpha; TNF-alpha; cartilage;  
KW ear; proliferation; glucose; free fatty acid; skeletal muscle; adipocyte;  
KW A-peptide; factor VIIA; gene therapy; ss.  
XX Homo sapiens.  
OS  
PN WO200140466-A2.  
XX  
PD 07-JUN-2001.  
XX  
PF 01-DEC-2000; 2000WO-US032678.  
XX  
PR 01-DEC-1999; 99WO-US028301.  
PR 01-DEC-1999; 99WO-US028634.  
PR 02-DEC-1999; 99WO-US028551.  
PR 02-DEC-1999; 99WO-US028584.  
PR 02-DEC-1999; 99WO-US028565.  
PR 09-DEC-1999; 99US-0170262P.  
PR 16-DEC-1999; 99WO-US030095.  
PR 20-DEC-1999; 99WO-US030911.  
PR 20-DEC-1999; 99WO-US030999.  
PR 30-DEC-1999; 99WO-US031243.  
PR 30-DEC-1999; 99WO-US031274.  
PR 05-JAN-2000; 2000WO-US000219.  
PR 06-JAN-2000; 2000WO-US000277.  
PR 06-JAN-2000; 2000WO-US00376.  
PR 11-FEB-2000; 2000WO-US003565.  
PR 18-FEB-2000; 2000WO-US004341.  
PR 18-FEB-2000; 2000WO-US004342.  
PR 22-FEB-2000; 2000WO-US004414.  
PR 24-FEB-2000; 2000WO-US004914.  
PR 24-FEB-2000; 2000WO-US005004.  
PR 01-MAR-2000; 2000WO-US005601.  
PR 02-MAR-2000; 2000WO-US005841.  
PR 03-MAR-2000; 2000US-0187202P.  
PR 10-MAR-2000; 2000WO-US006319.  
PR 15-MAR-2000; 2000WO-US006884.  
PR 20-MAR-2000; 2000WO-US007377.  
PR 21-MAR-2000; 2000WO-US007532.  
PR 30-MAR-2000; 2000WO-US008439.  
PR 17-MAY-2000; 2000WO-US013705.  
PR 22-MAY-2000; 2000WO-US014042.  
PR 30-MAY-2000; 2000WO-US014941.  
PR 02-JUN-2000; 2000WO-US015264.  
PR 05-JUN-2000; 2000US-0209832P.  
PR 28-JUL-2000; 2000WO-US020710.  
PR 11-AUG-2000; 2000WO-US022031.  
PR 23-AUG-2000; 2000WO-US023522.  
PR 24-AUG-2000; 2000WO-US023328.  
PR 08-NOV-2000; 2000WO-US030952.  
PR 10-NOV-2000; 2000WO-US030873.  
XX  
XX (GETH ) GENENTECH INC.  
XX  
XX Baker KP, Beresini M, Deforge L, Desnoyers L, Filvaroff E, Gao W;  
PI Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;  
PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;  
XX  
XX WPI; 2001-408281/43.  
DR P-PSDB; AAU12203.  
XX  
XX Isolated , secretory and transmembrane PRO polypeptide used to detect  
PT other PRO polypeptides, link bioactive molecules to cells expressing PRO  
PT polypeptides, and detect the presence of mammalian tumors e.g. lung,  
PT breast, prostate, cervical.  
XX  
XX Claim 3; Fig 63; 813pp; English.  
XX  
XX AAS21244-AAS21518 encode for novel human secretory and transmembrane PRO  
CC polypeptides. The PRO polypeptides are useful to detect other PRO  
CC

Db 641 GTGGTCCGGGAGCTGGTGGTACCTCCCTGGTGGAGTCAGCTCCAGGGGGTCCAGACCTAC 700  
 Qy 221 GluMetPheSerTrpPheLeuArgProLeuProGlyProHisGlyGlnLeuLeuHisHis 240  
 Db 701 GAGATGTTCTCTCTGGTCTCTGGGCCCCCTGGCAGGCCCCCAAGCAGCTCTCCACCAC 760  
 Qy 241 ValSerThrLeuAlaAlaPheThrValArgGlnValGlnGlnHisGlnGlyAsnLeuAsp 260  
 Db 761 GTCAGCAGCTTGGCTGCTTTCACAGTCCGGCAGGTGGCAGCAGCAGGGGAACCTGGAT 820  
 Qy 261 AlaSerGlyProAlaArgAspLeuValAspAlaPheLeuLeuLysMetAlaGlnGlu 280  
 Db 821 GCTTCGGGCCCCGACGAGCTTGTGATGCTTCTCTGCTGATGATGCAAGATGCAAGAGAA 880  
 Qy 281 GlnAsnProGlyThrGluPheThrAsnLysAsnMetLeuMetThrValIleTyrLeuLeu 300  
 Db 881 CAAAACCCAGGCACAGAAATTCACCAACAGAAACATGCTGATGACAGTCAITTAITTTGCTG 940  
 Qy 301 PheAlaGlyThrMetThrValSerThrThrValGlyTyrThrLeuLeuLeuMetLys 320  
 Db 941 TTTTCTGGGACGATGACGGTCAGCACACGCTCGGCTATACCTCTCTCTGATGAAA 1000  
 Qy 321 TyrProHisValGlnLysTrpValArgGlnGluLeuAsnArgGluLeuGlyAlaGlyGln 340  
 Db 1001 TACCTCATGTCCAAAGTGGTACGTGAGGAGCTGAATCGGAGCTGGGGCTGGCCAG 1060  
 Qy 341 AlaProSerLeuGlyAspArgThrArgLeuProTyrThrAspAlaValLeuHisGluAla 360  
 Db 1061 GCACCAAGCCTAGGGACCGTACCCGCTCTTACACCGACGGGTTCGATGAGCGC 1120  
 Qy 361 GlnArgLeuLeuAlaLeuValProMetGlyIleProArgThrLeuMetArgThrThrArg 380  
 Db 1121 CAGCGGCTGTGGGCTGTGGCCATGGAATACCCGACACCTCATGGGACCAACCCCGC 1180  
 Qy 381 PheArgGlyTyrThrLeuProGlnGlyThrGluValPheProLeuLeuGlySerIleLeu 400  
 Db 1181 TTCGAGGGTACACCTCGCCCGCAGGGCAGCGAGGTCTTCCCTCTCTTGGCTCCATCCTG 1240  
 Qy 401 HisAspProAsnIlePheLysHisProGluGluPheAsnProAspArgPheLeuAspAla 420  
 Db 1241 CATGACCCCAACATCTCAAGCACCCAGAGAGAGTTCAACCCAGACCCGTTTCTGGATGCA 1300  
 Qy 421 AspGlyArgPheArgLysHisGluAlaPheLeuProPheSerLeuGlyLysArgValCys 440  
 Db 1301 GATGGACGGTTCAGGAAGCATAGGCGTCTTCCCTCTCTTAGGAAGCGTGTCTGC 1360  
 Qy 441 LeuGlyGluGlyLeuAlaLysAlaGluLeuPheLeuPheThrThrIleLeuGlnAla 460  
 Db 1361 CTTGGAGAGCGCTGGCAAGCGGAGCTCTTCTCTTTCACCAACCATCTTACAGCC 1420  
 Qy 461 PheSerLeuGluSerProCysProProAspThrLeuSerLeuLysProThrValSerGly 480  
 Db 1421 TTTCTCTGGAGACCGTGGCCCGCGGACACCTGAGCCTCAAGCCCAACCGTCAGTGGC 1480  
 Qy 481 LeuPheAsnIleProProAlaPheGlnLeuGlnValArgProThrAspLeuHisSerThr 500  
 Db 1481 CTTTTCACATTCCTCCAGCTTCCAGTCCGAGTCCGTCGATCCGCTTCACTTCACTCCACC 1540  
 Qy 501 ThrGlnThrArg 504  
 Db 1541 ACGCAGACCAGA 1552  
 RESULT 8  
 ID ACD23884  
 XX ACD23884 standard; cDNA; 2623 BP.  
 XX  
 AC ACD23884;  
 XX  
 DT 26-AUG-2003 (first entry)  
 DE  
 DE Novel human secreted and transmembrane protein PRO1906 cDNA.  
 XX  
 XX Human; secreted and transmembrane protein; PRO; antiinflammatory;  
 KW

antiarteriosclerotic; cardiant; anti-infertility; anti-HIV; cytostatic;  
 antidiabetic; gene therapy; tumour necrosis factor (TNF)-alpha release;  
 TNF-alpha release; cell proliferation; cell differentiation;  
 gene expression modulator; proteoglycan release; cytokine release;  
 tumour; inflammatory disease; organ failure; atherosclerosis;  
 cardiac injury; infertility; birth defect; premature aging; AIDS;  
 acquired immunodeficiency syndrome; cancer; diabetic complication;  
 chromosome mapping; gene mapping; pharmaceutical; diagnostic; biosensor;  
 bioreactor; tissue typing; gene; ss.  
 Homo sapiens.  
 US2003032156-A1.  
 13-FEB-2003.  
 06-MAY-2002; 2002US-00140474.  
 31-MAR-1997; 97WO-US005230.  
 12-JUN-1998; 98WO-US012456.  
 14-JUL-1998; 98WO-US014552.  
 28-AUG-1998; 98WO-US017888.  
 10-SEP-1998; 98WO-US018824.  
 14-SEP-1998; 98WO-US019094.  
 14-SEP-1998; 98WO-US019177.  
 14-SEP-1998; 98WO-US019330.  
 17-SEP-1998; 98WO-US019437.  
 07-OCT-1998; 98WO-US021141.  
 29-OCT-1998; 98WO-US022991.  
 29-OCT-1998; 98WO-US022992.  
 01-DEC-1998; 98WO-US024855.  
 05-JAN-1999; 99WO-US000106.  
 08-MAR-1999; 99WO-US005028.  
 10-MAR-1999; 99WO-US005190.  
 20-APR-1999; 99WO-US008615.  
 14-MAY-1999; 99WO-US010733.  
 02-JUN-1999; 99WO-US012252.  
 01-SEP-1999; 99WO-US020111.  
 08-SEP-1999; 99WO-US020594.  
 13-SEP-1999; 99WO-US020944.  
 15-SEP-1999; 99WO-US021090.  
 15-SEP-1999; 99WO-US021547.  
 05-OCT-1999; 99WO-US023089.  
 29-NOV-1999; 99WO-US028214.  
 30-NOV-1999; 99WO-US028313.  
 30-NOV-1999; 99WO-US028409.  
 01-DEC-1999; 99WO-US028301.  
 01-DEC-1999; 99WO-US028634.  
 02-DEC-1999; 99WO-US028551.  
 02-DEC-1999; 99WO-US028564.  
 16-DEC-1999; 99WO-US028565.  
 20-DEC-1999; 99WO-US030095.  
 20-DEC-1999; 99WO-US030911.  
 22-DEC-1999; 99WO-US030999.  
 30-DEC-1999; 99WO-US030720.  
 30-DEC-1999; 99WO-US031243.  
 30-DEC-1999; 99WO-US031274.  
 05-JAN-2000; 2000WO-US000219.  
 06-JAN-2000; 2000WO-US000277.  
 11-FEB-2000; 2000WO-US003376.  
 18-FEB-2000; 2000WO-US003565.  
 18-FEB-2000; 2000WO-US004341.  
 22-FEB-2000; 2000WO-US004342.  
 24-FEB-2000; 2000WO-US004414.  
 24-FEB-2000; 2000WO-US004914.  
 01-MAR-2000; 2000WO-US005004.  
 02-MAR-2000; 2000WO-US005601.  
 02-MAR-2000; 2000WO-US005746.  
 10-MAR-2000; 2000WO-US005841.  
 15-MAR-2000; 2000WO-US006319.  
 2000WO-US006884.

CC to factor VIIA, or detect the presence of tumour in a mammal. The nucleic  
CC acid and polypeptide encoded by it, are useful for treating inflammatory  
CC diseases, organ failure, atherosclerosis, cardiac injury, infertility,  
CC birth defects, premature aging, acquired immunodeficiency syndrome  
CC (AIDS), cancer, or diabetic complications. The nucleic acid is useful as  
CC hybridisation probes, in chromosome and gene mapping, and in generating  
CC antisense RNA or DNA. The polypeptides are useful as pharmaceuticals,  
CC diagnostics, biosensors or bioreactors. Both are useful in tissue typing.  
CC This sequence encodes a novel human secreted and transmembrane PRO  
CC polypeptide  
XX  
SQ Sequence 2623 BP; 552 A; 865 C; 704 G; 502 T; 0 U; 0 Other;  
  
Alignment Scores:  
Pred. No.: 5, 37e-231 Length: 2623  
Score: 2615.00 Matches: 504  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 7 Gaps: 0  
  
US-10-669-693-2 (1-504) x ACD23884 (1-2623)  
  
QY 1 MetGluAlaThrGlyThrTrpAlaLeuLeuLeuAlaLeuAlaLeuLeuLeuLeuThr 20  
DB 41 ATGGAGGCGACCGGCACCTGCTGCTGGCGCTGGCGCTCTCTCTGCTGAGG 100  
QY 21 LeuAlaLeuSerGlyThrArgAlaArgGlyHisLeuProGlyProGlyProThrProLeuPro 40  
DB 101 CTGGCGCTGTCGGGACACCGAGGCCACCTGCGCCCGGGCCCGCCGCGGTACCA 160  
QY 41 LeuLeuGlyAsnLeuLeuGlnLeuArgProGlyAlaLeuLeuLeuLeuLeuMetArgLeu 60  
DB 161 CTGCTGGGAACTCTCTGAGCTACCGCCCGGGCGCTGCTATTAGGGCTCATGCGGCTG 220  
QY 61 SerLysLysTyrGlyProValPheThrIleTyrLeuGlyProThrArgProValValVal 80  
DB 221 AGTAAGAGTACGACCGGTGTTACCATCTACCTGGGACCTGCGGCGCTGTTGGTGGTC 280  
QY 81 LeuValGlyGlnGluAlaValArgGluAlaLeuGlyGlyGlnAlaGluGluPheSerGly 100  
DB 281 CTGGTTGGGCGAGGAGCTGTGGGAGGCCCTGGAGGCTGAGGCTGAGGAGTTCCAGGCG 340  
QY 101 ArgGlyThrValAlaMetLeuGluGlyThrPheAspGlyHisGlyValPhePheSerAsn 120  
DB 341 CGGGGAACTGATGATGCTGGAGGAGCTTTTGTATGGCATGGGCTTTTCTTCTCCAAC 400  
QY 121 GlyGluArgTrpArgGlnLeuArgLysPheThrMetLeuAlaLeuArgAspLeuGlyMet 140  
DB 401 GGGGAGCGGTGGAGGAGCTGAGGAAAGTTTACCATCTGCTCTGCGGGACCTGGGCGATG 460  
QY 141 GlyLysArgGluGlyGluGluLeuIleGlnAlaGluAlaArgCysLeuValGluThrPhe 160  
DB 461 GGGAGCGAGAGGCGGAGGAGCTGATCCAGCGGAGGCCCGGCTGTCTGGTGGAGACATTC 520  
QY 161 GlnGlyThrGluGlyArgProPheAspProSerLeuLeuLeuAlaGlnAlaThrSerAsn 180  
DB 521 CAGGGGACAGAGGAGCGCCATTCGATCCCTCTCTGCTGCGCCAGGCGACCTCCCAAC 580  
QY 181 ValValCysSerLeuLeuPheGlyLeuArgPheSerTyrIleGluAspLysGluPheGlnAla 200  
DB 581 GTAGTCTGCTCCCTCTCTTGGCCTCTTCTCTATGAGGATAGGAGTTCCAGGCC 640  
QY 201 ValValArgAlaAlaGlyGlyThrLeuLeuGlyValSerSerGlnGlyGlyGlnThrTyr 220  
DB 641 GTGCTCGGCGAGCTGTTGTGTACCTGCTGGAGTCACTCCCGAGGGGCTGAGACCTAC 700  
QY 221 GluMetPheSerTrpPheLeuArgProLeuProGlyProHisLysGlnLeuLeuHis 240  
DB 701 GAGATGTTCTCTGTTCTGCGGCCCTGCGGCGCCCTGCCAGGCGCCCAACAGAGCTCTCCAC 760  
QY 241 ValSerThrLeuAlaAlaPheThrValArgGlnValGlnGlnHisGlnGlyAsnLeuAsp 260

PR 20-MAR-2000; 2000WO-US007377.  
PR 21-MAR-2000; 2000WO-US007532.  
PR 30-MAR-2000; 2000WO-US008439.  
PR 17-MAY-2000; 2000WO-US013705.  
PR 22-MAY-2000; 2000WO-US014042.  
PR 30-MAY-2000; 2000WO-US014941.  
PR 02-JUN-2000; 2000WO-US015264.  
PR 28-JUL-2000; 2000WO-US020710.  
PR 11-AUG-2000; 2000WO-US022031.  
PR 23-AUG-2000; 2000WO-US023352.  
PR 24-AUG-2000; 2000WO-US023328.  
PR 08-NOV-2000; 2000WO-US030952.  
PR 10-NOV-2000; 2000WO-US030873.  
PR 01-DEC-2000; 2000WO-US032678.  
PR 20-DEC-2000; 2000US-00747259.  
PR 20-DEC-2000; 2000WO-US034956.  
PR 28-FEB-2001; 2001US-00796498.  
PR 28-FEB-2001; 2001WO-US006520.  
PR 01-MAR-2001; 2001WO-US006666.  
PR 09-MAR-2001; 2001US-00802706.  
PR 14-MAR-2001; 2001US-00808689.  
PR 22-MAR-2001; 2001US-00816744.  
PR 05-APR-2001; 2001US-00828366.  
PR 10-MAY-2001; 2001US-00854280.  
PR 10-MAY-2001; 2001US-00854280.  
PR 18-MAY-2001; 2001US-00860216.  
PR 25-MAY-2001; 2001US-00866028.  
PR 25-MAY-2001; 2001US-00866034.  
PR 25-MAY-2001; 2001WO-US017092.  
PR 01-JUN-2001; 2001US-00872035.  
PR 01-JUN-2001; 2001WO-US017800.  
PR 05-JUN-2001; 2001US-00874503.  
PR 14-JUN-2001; 2001US-00882636.  
PR 19-JUN-2001; 2001US-00886342.  
PR 20-JUN-2001; 2001WO-US019692.  
PR 21-JUN-2001; 2001US-00887879.  
PR 22-JUN-2001; 2001WO-US020116.  
PR 29-JUN-2001; 2001WO-US021066.  
PR 09-JUL-2001; 2001WO-US021735.  
PR 18-JUL-2001; 2001US-00908827.  
PR 06-AUG-2001; 2001US-00924419.  
PR 09-AUG-2001; 2001US-00927796.  
PR 16-AUG-2001; 2001US-00931836.  
PR 19-DEC-2001; 2001US-00028072.  
  
XX (GETH ) GENENTECH INC.  
XX Baker KP, Beresini M, Deforge L, Deenoyers L, Filvaroff E, Gao W;  
PI Gerritsen MB, Goddard A, Godowski PJ, Gurney AL, Sherwood S;  
PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;  
  
XX WPI: 2003-341980/32.  
DR P-PSDB; ABO17647.  
DR  
  
XX New secreted and transmembrane PRO nucleic acids, for treating  
PT inflammation, organ failure, atherosclerosis, cardiac injury,  
PT infertility, birth defects, premature aging, acquired immunodeficiency  
PT syndrome (AIDS), or cancer.  
XX  
PS Claim 2; Fig 63; 660pp; English.  
XX  
XX The invention describes an isolated nucleic acid (I) comprising, or which  
CC has 80 % sequence identity to, or the full-length coding sequence of, one  
CC of 275 nucleotide sequences, and which encodes a corresponding  
CC polypeptide selected from 275 amino acid sequences, where all sequences  
CC are given in the specification. The polypeptide encoded by (I) is used to  
CC detect PRO polypeptides, link a bioactive molecule to a cell expressing a  
CC PRO polypeptide, modulate a biological activity of a cell, stimulate the  
CC release of tumour necrosis factor (TNF)-alpha from human blood, modulate  
CC the uptake of glucose or free fatty acid by cells, stimulate or inhibit  
CC the proliferation or differentiation of cells or gene expression,  
CC stimulate the release of proteoglycans, stimulate the release of cytokine  
CC from peripheral blood mononuclear cells, inhibit the binding of A-peptide



```
Db 761 GTCAGCACCTTGGCTGCTTCCAGTCGCGCAGGTGCAGCAGCAGCGGAACCTGGAT 820
Qy 261 AlaSerGlyProAlaAspLeuValAspAlaPheLeuLeuLysMetAlaGlnGlu 280
Db 821 GCTTCGGGCCCCGACGTGACCTTGTGATGCCCTTCCTGCTGGAAGATGGCAGAGGAA 880
Qy 281 GlnAsnProGlyThrGluPheThrAsnLysAsnMetLeuMetThrValIleTyrLeuLeu 300
Db 881 CAAAACCCAGGCACAGAATTCCACCAACAAGACATGCTGATGACAGTCAATTAATTGCTG 940
Qy 301 PheAlaGlyThrMetThrValSerThrThrValGlyTyrThrLeuLeuLeuMetLys 320
Db 941 TTTCTGCGGACGATGACGGTCAGCACCGCTGCGCTATACCTCTCTCTCTGATGAA 1000
Qy 321 TyrProHisValGlnLysTyrValArgGluGluLeuAsnArgGluLeuGlyAlaGlyGln 340
Db 1001 TACCTCATGTCCTCAAAAGTGGTACGTGAGGAGCTGAATCGGAGCTGGGGGCTGGCCAG 1060
Qy 341 AlaProSerLeuGlyAspArgThrArgLeuProTyrThrAspAlaValLeuHisGluAla 360
Db 1061 GCACCAAGCCTAGGGACCGTACCGGCTCCCTTACACCGACGCGGTCTGTCATGAGCG 1120
Qy 361 GlnArgLeuLeuAlaLeuValProMetGlyIleProArgThrLeuMetArgThrThrArg 380
Db 1121 CAGCGGCTGCTGGCGCTGGTGCCTATGGGAATACCCGACACCTCATGCGGACCAACCCGC 1180
Qy 381 PheArgGlyTyrThrLeuProGlnGlyThrGluValPheProLeuLeuGlySerIleLeu 400
Db 1181 TTCGAGGTTACACCTTGCCCGAGGACGAGGTCTTCCCTCTCTTGGCTCCATCCTG 1240
Qy 401 HisAspProAsnIlePheLysHisProGluGluPheAsnProAspArgPheLeuAspAla 420
Db 1241 CATGACCCCAACATCTTCAAGCACCCAGAGAGTTCACCCAGACCGTTTCTGGATGCA 1300
Qy 421 AspGlyArgPheArgLysHisGluAlaPheLeuProPheSerLeuGlyLysArgValCys 440
Db 1301 GATGGACGGTTACGAGCATGAGGCTTCTGCTCTCTCTTAGGGAAGCGTGTCTGC 1360
Qy 441 LeuGlyGluGlyLeuAlaLysAlaGluPheLeuPheLeuPheThrIleLeuGlnAla 460
Db 1361 CTTGGAGAGGCTTGGAAAGCGGAGCTCTTCTCTTCTTCCACCATCTTACAGGC 1420
Qy 461 PheSerLeuGluSerProCysProAspThrLeuSerLeuLysProThrValSerGly 480
Db 1421 TTCTCTGCTGAGAGCGCTGCCCGCGCACCTGAGCTCAAGCCACCGCTAGTGGC 1480
Qy 481 LeuPheAsnIleProProAlaPheGlnLeuGlnValArgProThrAspLeuHisSerThr 500
Db 1481 CTTTTCACATTCCTCCAGCTTCCAGCTGCAAGTCCGTCCTCCACTGACCTTCACTCCACC 1540
Qy 501 ThrGlnThrArg 504
Db 1541 ACGCAGACCCAGA 1552

RESULT 9
ACA67025
ID ACA67025 standard; cDNA; 2623 BP.
XX
AC ACA67025;
XX
DT 23-JUN-2003 (first entry)
XX
DE cDNA encoding human PRO polypeptide #32.
XX
KW Human; PRO polypeptide; secreted and transmembrane protein;
anti-PRO antibody; diagnostic assay; gene expression; diabetes;
bone disorder; cartilage disorder; rheumatoid arthritis; obesity;
KW sports injury; osteoarthritis; hyper-insulinaemia; hypo-insulinaemia;
KW hearing loss; coagulation disorder; stroke; heart attack; cardiac;
KW antidiabetic; anorectic; vulnery; antiarthritic; osteopathic;
KW antirheumatic; auditory; cerebroprotective; angiogenic; gene; ss.
XX
OS Homo sapiens.
```

```
XX US20030043111-Al.
XX 02-JAN-2003.
XX 19-DEC-2001; 2001US-00028072.
XX 18-JUN-1997; 97US-0049911P.
XX 26-AUG-1997; 97US-0056974P.
XX 17-SEP-1997; 97US-0059113P.
XX 17-SEP-1997; 97US-0059115P.
XX 17-SEP-1997; 97US-0059117P.
XX 17-SEP-1997; 97US-0059122P.
XX 17-SEP-1997; 97US-0059184P.
XX 18-SEP-1997; 97US-0059263P.
XX 19-SEP-1997; 97US-0059352P.
XX 19-SEP-1997; 97US-0059588P.
XX 24-SEP-1997; 97US-0059836P.
XX 17-OCT-1997; 97US-0062250P.
XX 17-OCT-1997; 97US-0062285P.
XX 17-OCT-1997; 97US-0062287P.
XX 17-OCT-1997; 97US-0063755P.
XX 24-OCT-1997; 97US-0062814P.
XX 24-OCT-1997; 97US-0062816P.
XX 24-OCT-1997; 97US-0063045P.
XX 24-OCT-1997; 97US-0063082P.
XX 24-OCT-1997; 97US-0063127P.
XX 27-OCT-1997; 97US-0063327P.
XX 27-OCT-1997; 97US-0063329P.
XX 28-OCT-1997; 97US-0063550P.
XX 28-OCT-1997; 97US-0063561P.
XX 29-OCT-1997; 97US-0063704P.
XX 29-OCT-1997; 97US-0063733P.
XX 29-OCT-1997; 97US-0063735P.
XX 29-OCT-1997; 97US-0063738P.
XX 03-NOV-1997; 97US-0064488P.
XX 07-NOV-1997; 97US-0064809P.
XX 12-NOV-1997; 97US-0065186P.
XX 17-NOV-1997; 97US-0065846P.
XX 21-NOV-1997; 97US-0066364P.
XX 24-NOV-1997; 97US-0066453P.
XX 24-NOV-1997; 97US-0066511P.
XX 24-NOV-1997; 97US-0066770P.
XX 11-DEC-1997; 97US-0069212P.
XX 11-DEC-1997; 97US-0069278P.
XX 11-DEC-1997; 97US-0069334P.
XX 16-DEC-1997; 97US-0069694P.
XX 23-JAN-1998; 98US-0072320P.
XX 04-FEB-1998; 98US-0073612P.
XX 09-FEB-1998; 98US-0074086P.
XX 09-FEB-1998; 98US-0074092P.
XX 12-MAR-1998; 98US-0077791P.
XX 20-MAR-1998; 98US-0078910P.
XX 25-MAR-1998; 98US-0079294P.
XX 27-MAR-1998; 98US-0079663P.
XX 27-MAR-1998; 98US-0079728P.
XX 31-MAR-1998; 98US-0080165P.
XX 12-JUN-1998; 98WO-US012456.
XX 14-JUL-1998; 98WO-US014552.
XX 28-AUG-1998; 98WO-US017888.
XX 10-SEP-1998; 98WO-US018824.
XX 14-SEP-1998; 98WO-US019093.
XX 14-SEP-1998; 98WO-US019094.
XX 14-SEP-1998; 98WO-US019177.
XX 16-SEP-1998; 98WO-US019330.
XX 17-SEP-1998; 98WO-US019437.
XX 07-OCT-1998; 98WO-US021141.
XX 29-OCT-1998; 98WO-US022991.
XX 29-OCT-1998; 98WO-US022992.
XX 20-NOV-1998; 98WO-US024855.
XX 01-DEC-1998; 98WO-US025108.
XX 05-JAN-1999; 99WO-US000106.
XX 08-MAR-1999; 99WO-US005028.
```

PR	10-MAR-1999;	99WO-US005190.	DB:	7	Gaps:	0
PR	20-APR-1999;	99WO-US008615.	US-10-669-693-2 (1-504) x ACA67025 (1-2623)			
PR	14-MAY-1999;	99WO-US010733.	Qy	1	MetGluAlaThrGlyThrTrpAlaLeuLeuLeuAlaLeuAlaLeuLeuLeuThr	20
PR	02-JUN-1999;	99WO-US012252.	Db	41	ATGGAGCGACCGGCACCTGGCGCTGCTGTGGCGCTGGCGCTCTCTCTGTGTGTGACG	100
PR	01-SEP-1999;	99WO-US020111.	Qy	21	LeuAlaLeuSerGlyThrArgAlaArgGlyHisLeuProGlyProGlyProLeuPro	40
PR	08-SEP-1999;	99WO-US020594.	Db	101	CTGGCGCTGTCGGGACCGAGGCGCCGAGGCGCACCTGCCCGCCCGGCGCCACCGCGCTACCA	160
PR	13-SEP-1999;	99WO-US020944.	Qy	41	LeuLeuGlyAsnLeuLeuGlnLeuArgProGlyAlaLeuLeuLeuSerGlyLeuMetArgLeu	60
PR	15-SEP-1999;	99WO-US021090.	Db	161	CTGCTGGGAAACCTCTCTCAGCTACGCGCGCGGCGCTGTTATTCAGGCGCTCATGGCGGTG	220
PR	05-OCT-1999;	99WO-US021547.	Qy	61	SerLysLysTyrGlyProValPheThrIleTyrLeuGlyProTrpArgProValValVal	80
PR	29-NOV-1999;	99WO-US028214.	Db	221	AGTAAGAAGTACGGACCGGTGTTACCATCTACCTGGGACCGCTGGCGGCTGTGGTGGTC	280
PR	30-NOV-1999;	99WO-US028313.	Qy	81	LeuValGlyGlnGluAlaValArgGluAlaLeuGlyGlyGlnAlaGluGluPheSerGly	100
PR	01-DEC-1999;	99WO-US028301.	Db	281	CTGGTTGGCGAGAGGCTGTGGGAGGCGCTGGGAGGTGAGCTGAGGAGTTCAGCGGC	340
PR	02-DEC-1999;	99WO-US028551.	Qy	101	ArgGlyThrValAlaMetLeuGlyThrPheAspGlyHisGlyValPhePheSerAsn	120
PR	02-DEC-1999;	99WO-US028564.	Db	341	CGGGGAAACCGTAGCGATGCTGGAGGGACTTTTGATGGCCATGGGGTTTTCTTCTCCAAC	400
PR	02-DEC-1999;	99WO-US028585.	Qy	121	GlyGluArgTrpArgGlnLeuArgLysPheThrMetLeuAlaLeuArgAspLeuGlyMet	140
PR	16-DEC-1999;	99WO-US030095.	Db	401	GGGAGCGGTGGAGGCGAGCTGTCGCGGAGGCGCTGGGAGGTGCTGCTGCGGACCTGGGCATG	460
PR	20-DEC-1999;	99WO-US030911.	Qy	141	GlyLysArgGluGlyGluLeuLeuLeuGlnAlaGluAlaArgCysLeuValGluThrPhe	160
PR	30-DEC-1999;	99WO-US031274.	Db	461	GGGAAGCGAGAGAGGCGAGAGCTGATCCAGGGCGAGGCGCGGTGCTGTGGAGACATTC	520
PR	05-JAN-2000;	2000WO-US000219.	Qy	161	GlnGlyThrGluGlyArgProPheAspProSerLeuLeuLeuAlaGlnAlaThrSerAsn	180
PR	06-JAN-2000;	2000WO-US000277.	Db	521	CAGGGGACAGAGGACGCCCATTCGATCCCTCCCTGCTGCTGGCCCGAGGCGCTCTCCAAC	580
PR	11-FEB-2000;	2000WO-US000376.	Qy	181	ValValCysSerLeuLeuPheGlyLeuArgPheSerTyrGluAspLysGluPheGlnAla	200
PR	18-FEB-2000;	2000WO-US004341.	Db	581	GTAGTCTGCTCCCTCTCTTGGCCCTCCGCTTCTCTATGAGGATGAAGAGTTCCAGGCC	640
PR	22-FEB-2000;	2000WO-US004414.	Qy	201	ValValArgAlaAlaGlyGlyThrLeuLeuGlyValSerSerGlnGlyGlyGlnThrTyr	220
PR	24-FEB-2000;	2000WO-US004914.	Db	641	GTGGTCCGGGCGAGCTGGTGTACCTGCTGGGAGTCCAGCTCCCGGGGGTCCAGACCTAC	700
PR	24-FEB-2000;	2000WO-US005004.	Qy	221	GluMetPheSerTrpPheLeuArgProLeuProGlyProHisLysGlnLeuLeuHisHis	240
PR	01-MAR-2000;	2000WO-US005601.	Db	701	GAGATGTTCTCTGGTTCTGGGCGCCCTGCCAGGCGCCCGCCACCAAGCAGCTCTCTCCACAC	760
PR	02-MAR-2000;	2000WO-US005746.	Qy	241	ValSerThrLeuAlaAlaPheThrValArgGlnValGlnGlnHisGlnGlyAsnLeuAsp	260
XX	WPI: 2003-352836/33.		Db	761	GTGAGCACCTTGGCTGCTTCAAGTCCGCGAGGTGCGAGCAGCAGCAGGGAACCTGGAT	820
XX	P-FSDB; ABU80901.		Qy	261	AlaSerGlyProAlaArgAspLeuValAspAlaPheLeuLeuLysMetAlaGlnGluGlu	280
XX	New isolated PRO polypeptide useful for treating diabetes, rheumatoid		Db	821	GCTTCGGGCCCCACAGTGCATCTTCTCGATGCTTCTGCTGAGATGGCCACAGGAGAA	880
XX	arthritis, sports injuries, obesity, hearing loss in mammals, stroke, or		Qy	281	GlnAsnProGlyThrGluPheThrAsnLysAsnMetLeuMetThrValIleTyrLeuLeu	300
XX	heart attack.		Db	881	CAAAACCCAGGCACAGAAATTCACCAAGAACATGCTGATGACAGTCAATTTATTTGCTG	940
XX	Claim 2; Fig 63; 643pp; English.		Qy	301	PheAlaGlyThrMetThrValSerThrThrValGlyTyrThrLeuLeuLeuMetLys	320
XX	The present invention relates to the isolation of novel human PRO		Db	941	TTTGCTGGGACGATGACGTCAGCACCGCTGGGCTATACCTCTCTGCTCTCTGATGAAA	1000
XX	polypeptides, and the polynucleotide sequences encoding them. The PRO		Qy	321	TyrProHisValGlnLysTrpValArgGluGluLeuAsnArgGluLeuGlyAlaGlyGln	340
XX	polypeptides are secreted and transmembrane proteins. The PRO		Db	1001	TACCTCATGTCACCAAGTGGTACGTGAGGAGCTGAATCGGGAGCTGGGGCTGGCGAG	1060
XX	polypeptides and polynucleotides are useful for preparing a medicament		Qy	341	AlaProSerLeuGlyAspArgThrArgLeuProTyrThrAspAlaValLeuHisGluAla	360
XX	useful in the treatment of diabetes, bone and/or cartilage disorders					
XX	(e.g. rheumatoid arthritis, sports injuries, osteoarthritis), obesity,					
XX	hyper- or hypo-insulinaemia, hearing loss, and coagulation disorders					
XX	(e.g. stroke, heart attack). Anti-PRO antibodies are useful in diagnostic					
XX	assays for PRO, by detecting its expression in specific cells, tissues or					
XX	serum, and for affinity purification of PRO from recombinant cell culture					
XX	or natural sources. ACA65994-ACA67268 represent cDNA sequences encoding					
XX	the human PRO polypeptides of the invention. Note: The sequence data for					
XX	this patent was obtained in electronic format directly from the USPTO web					
XX	site at seqdata.uspto.gov/psipsoIDEntry.html					
XX	Sequence 2623 BP; 552 A; 865 C; 704 G; 502 T; 0 U; 0 Other;					
XX	Alignment Scores:					
XX	Pred. No.:	5.37e-231	Length:	2623		
XX	Score:	2615.00	Matches:	504		
XX	Percent Similarity:	100.00%	Conservative:	0		
XX	Best Local Similarity:	100.00%	Mismatches:	0		
XX	Query Match:	100.00%	Indels:	0		

```

Db      1061 GCACCAAGCCTAGGGAGCGTACCCGCTTCCCTTACCGAGCGGTTCTGATGAGCGG 1120
Qy      361 GlnArgLeuLeuAlaLeuValProMetGlyLeProArgThrLeuMetArgThrThrArg 380
Db      1121 CAGCGGCTGCTGGCGCTGGTGGCCATGGGAATACCCGACCCCTCATGGGACCAACCCGC 1180
Qy      381 PheArgGlyTyrThrLeuProGlnGlyThrGluValPheProLeuLeuGlySerIleLeu 400
Db      1181 TTCGAGGGTACACCTGCCCGAGGCGACGGAGTCTTCCCCCTCTTGGCTCCATCCTG 1240
Qy      401 HisAspProAsnIlePheLysHisProGluGluPheAsnProAspArgPheLeuAspAla 420
Db      1241 CATGACCCCAACATCTTCAAGCACCACAGAGATTCAACCCAGACCGTTTCTCGATGCA 1300
Qy      421 AspGlyArgPheArgLysHisGluAlaPheLeuProPheSerLeuGlyLysArgValCys 440
Db      1301 GATGACCGTTTCAGAGACATGAGCGTTCCTGCGCTTCTCTAGGAAGCGTGTCTGC 1360
Qy      441 LeuGlyGluGlyLeuAlaLysAlaGluLeuPheLeuPheThrThrIleLeuGlnAla 460
Db      1361 CTTGAGAGGCGCTGGCAAGAGGAGCTTCTTCTTCTTACACCATCTTCAAGCC 1420
Qy      461 PheSerLeuGluSerProCysProAspThrLeuSerLeuLysProThrValSerGly 480
Db      1421 TTCTCCCTGGAGAGCGCGTGGCCGCGGACACCTGAGCCTCAAGCCACCGTCAGTGC 1480
Qy      481 LeuPheAsnIleProProAlaPheGlnLeuGlnValArgProThrAspLeuHisSerThr 500
Db      1481 CTTTTCACATTTCCCGAGCCTTCCAGTCCGAGTCCGCTCCACTGACCTTCACTCCACC 1540
Qy      501 ThrGlnThrArg 504
Db      1541 ACGCAGACGAGA 1552

RESULT 10
ACA03634
ID ACA03634 standard; cDNA; 2623 BP.
XX
AC ACA03634;
XX
DT 23-MAY-2003 (first entry)
XX
DE cDNA encoding human PRO polypeptide #32.
XX
KW Human; PRO polypeptide; secreted and transmembrane protein;
KW tumour necrosis factor-alpha; TNF-alpha; blood; proliferation;
KW differentiation; chondrocyte; tumour; genetic disorder; cytostatic; gene;
KW SS.
XX
OS Homo sapiens.
XX
PN US2003036180-A1.
XX
PD 20-FEB-2003.
XX
PF 09-MAY-2002; 2002US-00143114.
XX
PR 31-MAR-1997; 97WO-US005230.
PR 12-JUN-1998; 98WO-US012456.
PR 14-JUL-1998; 98WO-US014552.
PR 28-AUG-1998; 98WO-US017888.
PR 10-SEP-1998; 98WO-US018824.
PR 14-SEP-1998; 98WO-US015093.
PR 14-SEP-1998; 98WO-US019094.
PR 14-SEP-1998; 98WO-US019177.
PR 16-SEP-1998; 98WO-US019330.
PR 17-SEP-1998; 98WO-US019437.
PR 07-OCT-1998; 98WO-US021141.
PR 29-OCT-1998; 98WO-US022991.
PR 29-OCT-1998; 98WO-US022992.
PR 20-NOV-1998; 98WO-US024855.
PR 01-DEC-1998; 98WO-US025108.
PR 05-JAN-1999; 99WO-US000106.
PR 08-MAR-1999; 99WO-US005028.
PR 10-MAR-1999; 99WO-US005190.
PR 20-APR-1999; 99WO-US008615.
PR 14-MAY-1999; 99WO-US010733.
PR 02-JUN-1999; 99WO-US012252.
PR 01-SEP-1999; 99WO-US020111.
PR 08-SEP-1999; 99WO-US020594.
PR 13-SEP-1999; 99WO-US020944.
PR 15-SEP-1999; 99WO-US021090.
PR 15-SEP-1999; 99WO-US021547.
PR 05-OCT-1999; 99WO-US023089.
PR 29-NOV-1999; 99WO-US028214.
PR 30-NOV-1999; 99WO-US028313.
PR 30-NOV-1999; 99WO-US028409.
PR 01-DEC-1999; 99WO-US028301.
PR 01-DEC-1999; 99WO-US028634.
PR 02-DEC-1999; 99WO-US028551.
PR 02-DEC-1999; 99WO-US028564.
PR 02-DEC-1999; 99WO-US028565.
PR 16-DEC-1999; 99WO-US030095.
PR 20-DEC-1999; 99WO-US030911.
PR 20-DEC-1999; 99WO-US030999.
PR 22-DEC-1999; 99WO-US030720.
PR 30-DEC-1999; 99WO-US031243.
PR 30-DEC-1999; 99WO-US031274.
PR 05-JAN-2000; 2000WO-US000219.
PR 06-JAN-2000; 2000WO-US000277.
PR 06-JAN-2000; 2000WO-US000376.
PR 11-FEB-2000; 2000WO-US000356.
PR 18-FEB-2000; 2000WO-US004341.
PR 18-FEB-2000; 2000WO-US004342.
PR 22-FEB-2000; 2000WO-US004414.
PR 24-FEB-2000; 2000WO-US004914.
PR 01-MAR-2000; 2000WO-US005004.
PR 02-MAR-2000; 2000WO-US005746.
PR 02-MAR-2000; 2000WO-US005841.
PR 10-MAR-2000; 2000WO-US006319.
PR 15-MAR-2000; 2000WO-US006884.
PR 20-MAR-2000; 2000WO-US007377.
PR 21-MAR-2000; 2000WO-US007532.
PR 30-MAR-2000; 2000WO-US008439.
PR 17-MAY-2000; 2000WO-US013705.
PR 22-MAY-2000; 2000WO-US014042.
PR 30-MAY-2000; 2000WO-US014941.
PR 02-JUN-2000; 2000WO-US015264.
PR 28-JUL-2000; 2000WO-US020710.
PR 11-AUG-2000; 2000WO-US022031.
PR 23-AUG-2000; 2000WO-US023522.
PR 24-AUG-2000; 2000WO-US023328.
PR 08-NOV-2000; 2000WO-US030952.
PR 10-NOV-2000; 2000WO-US030873.
PR 01-DEC-2000; 2000WO-US032678.
PR 20-DEC-2000; 2000US-00747259.
PR 20-DEC-2000; 2000WO-US034956.
PR 28-FEB-2001; 2001US-00796498.
PR 28-FEB-2001; 2001WO-US006520.
PR 01-MAR-2001; 2001WO-US006666.
PR 09-MAR-2001; 2001US-00802706.
PR 14-MAR-2001; 2001US-00808689.
PR 22-MAR-2001; 2001US-00816744.
PR 05-APR-2001; 2001US-00828366.
PR 10-MAY-2001; 2001US-00854208.
PR 10-MAY-2001; 2001US-00854280.
PR 18-MAY-2001; 2001US-00860216.
PR 25-MAY-2001; 2001US-00866028.
PR 25-MAY-2001; 2001US-00866034.
PR 25-MAY-2001; 2001WO-US017092.
PR 01-JUN-2001; 2001US-00872035.
PR 01-JUN-2001; 2001WO-US017800.
PR 05-JUN-2001; 2001US-00874503.
PR 14-JUN-2001; 2001US-00882636.

```

19-JUN-2001; 2001US-00886342.  
 20-JUN-2001; 2001WO-US019692.  
 21-JUN-2001; 2001US-00887879.  
 22-JUN-2001; 2001WO-US020116.  
 29-JUN-2001; 2001WO-US021066.  
 09-JUL-2001; 2001WO-US021735.  
 18-JUL-2001; 2001US-00908827.  
 06-AUG-2001; 2001US-00924419.  
 09-AUG-2001; 2001US-00927796.  
 16-AUG-2001; 2001US-00931836.  
 19-DEC-2001; 2001US-00028072.  
 (GETH ) GENENTECH INC.  
 Baker KP, Beresini M, Deforge L, Desnoyers L, Filvaroff E, Gao W;  
 Gerlitsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;  
 Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WT, Zhang Z;  
 WPI; 2003-332040/31.  
 P-PSDB; ABU66601.  
 New secreted and transmembrane PRO nucleic acids, useful for gene  
 therapy, in chromosome and gene mapping, as chromosome markers, in tissue  
 typing, and in chromosome identification.  
 Claim 2; Fig 63; 660pp; English.  
 The present invention relates to the isolation of novel human PRO  
 polypeptides, and the polynucleotide sequences encoding them. The PRO  
 polypeptides are secreted and transmembrane proteins. The PRO  
 polypeptides are useful for detecting other PRO polypeptides, for linking  
 bioactive molecules to cells expressing PRO polypeptides, for modulating  
 biological activities of cells expressing PRO polypeptides, and for for  
 identifying agonists or antagonists. The PRO polypeptides are useful for  
 stimulating the release of tumour necrosis factor (TNF)-alpha from  
 human blood, for stimulating the proliferation or differentiation of  
 chondrocytes, and detecting the presence of tumours. The polynucleotide  
 sequences encoding PRO polypeptides are useful as hybridisation probes,  
 in chromosome and gene mapping, in the generation of antisense RNA and  
 DNA, in the preparation of PRO polypeptides, for generating transgenic  
 animals or knockout animals, for the genetic analysis of individuals with  
 genetic disorders, and in gene therapy. ACA03603-ACA03877 represent cDNAs  
 encoding the human PRO polypeptides of the invention. Note: The sequence  
 data for this patent was obtained in electronic format directly from the  
 USPTO web site at [seqdata.uspto.gov/psipdsIDentry.html](http://seqdata.uspto.gov/psipdsIDentry.html)  
 Sequence 2623 BP; 552 A; 865 C; 704 G; 502 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 5,37e-231 Length: 2623  
 Score: 2615.00 Matches: 504  
 Percent Similarity: 100.00% Conservatives: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 7 Gaps: 0

US-10-669-693-2 (1-504) x ACA03634 (1-2623)

1 MetGluAlaThrGlyThrTrpAlaLeuLeuLeuAlaLeuLeuLeuLeuLeuThr 20  
 41 ATGAGGCGACCGGACCTGGCGCTGCTGCGCGTGGCGTGTCTCTCTGCTGCTGACG 100  
 21 LeuAlaLeuSerGlyThrArgAlaArgGlyHisLeuProProGlyProThrProLeuPro 40  
 101 CTGGCGTGTCTCGGACCGAGGCGACCTGCGCGCGCGGCGCGCGCGCTACCA 160  
 41 LeuLeuGlyAsnLeuLeuGlnLeuArgProGlyValAlaLeuTyrsGlyLeuMetArgLeu 60  
 161 CTGCTGGGAACCTCTCTGAGCTACGCGCGCGGCGCTGTATTACGGGCTCATCGCGCTG 220  
 61 SerLysLysTyrsGlyProValPheThrIleTyrsLeuGlyProThrProValValVal 80  
 221 AGTAAGAAGTACGGACCGGTGTTTACCATCTACCTGGGACCGCTGGCGCGCTGTGTGTC 280

81 LeuValGlyGlnGluAlaValArgGluAlaLeuGlyGlyGlnAlaGluGluPheSerGly 100  
 281 CTGGTTGGGAGGAGGCTGTGGGAGGCCCTGGAGGTGAGGTGAGGTGAGGTGAGGTGAG 340  
 101 ArgGlyThrValAlaMetLeuGluGlyThrPheAspGlyHisGlyValPhePheSerAsn 120  
 341 CGGGAAACCGTAGCATGTCTGAAAGGACATTTTATGGCCATGGGGTTTTCTTCTCCAAAC 400  
 121 GlyGluArgTrpArgGlnLeuArgLysPheThrMetLeuAlaLeuArgAspLeuGlyMet 140  
 401 GGGAGCGGTGGAGCAGCTGAGGAAGTTTACCATGCTTGTCTCTGGGAGCCTGGGCGATG 460  
 141 GlyLysArgGluGlyGluGluLeuGlnAlaGluAlaArgCysLeuValGluThrPhe 160  
 461 GGGAGCGAGGAAGCGAGAGCTGATCAGCGGAGGCCCGGTGTCTGTGTGAGACATTC 520  
 161 GlnGlyThrGluGlyArgProPheAspProSerLeuLeuAlaGlnAlaThrSerAsn 180  
 521 CAGGGGACAGAAGGACGCCCATTCGATCCTCCCTGCTGCTGGCCAGGCCACCTCCAAAC 580  
 181 ValValCysSerLeuLeuPheGlyLeuArgPheSerTyrsGluAspLysGluPheGlnAla 200  
 581 GTAGTCTGCTCCTCCTCTTTGGCCCTCGCTTCTCTATGAGGATAAGGAGTTCCAGGCC 640  
 201 ValValArgAlaAlaGlyGlyThrLeuLeuGlyValSerSerGlnGlyGlyGlnThrTyr 220  
 641 GTGGTCCGGGAGCTGGTGGTACCTGCTGGGAGTCAGCTCCAGGGGGGTGAGACCTAC 700  
 221 GluMetPheSerTrpPheLeuArgProGlyProGlyProHisLysGlnLeuLeuHis 240  
 701 GAGATGTTCTCTCTGCTGCTCTGGGCGCTTCCAGGCCCGCCCAAGCAGCTCTCTCCACCAC 760  
 241 ValSerThrLeuAlaAlaPheThrValArgGlnValGlnGlnHisGlnGlyAsnLeuAsp 260  
 761 GTGAGCCTTGGCTGCTCTTCAAGTCCGGGAGGTGCAGCAGCAGCAGGAGAACCTGGAT 820  
 261 AlaSerGlyProAlaArgAspLeuValAspAlaPheLeuLeuLysMetAlaGlnGluGlu 280  
 821 GCTTCGGGCGCCGCGAGCTGACCTTGTGGATGCTTCTCTGCTGAAGATGGCACAGGAGAA 880  
 281 GlnAsnProGlyThrGluPheThrAsnLysAsnMetLeuMetThrValIleTyrsLeuLeu 300  
 881 CAAACCCAGGACAGAAATTCACCAAGAACATGCTGATGACAGTCAATTTATTGTCTG 940  
 301 PheAlaGlyThrMetThrValSerThrValGlyTyrThrLeuLeuLeuLeuMetLys 320  
 941 TTTGCTGGGACGATGACGCTCAGCACCGGTGGCTATACCTCTCTCTCTGATGATAA 1000  
 321 TyrProHisValGlnLysTrpValArgGluGluLeuAsnArgGluLeuGlyAlaGlyGln 340  
 1001 TACCTCATGTCCAAAAGTGGGTACGTGAGGAGCTGAATCGGGAGCTGGGGGCTGGCCAG 1060  
 341 AlaProSerLeuGlyAspArgThrArgLeuProTyThrAspAlaValLeuHisGluAla 360  
 1061 GCACCAAGCCTAGGGGACCGTACCCGCTCCCTTACACCGACCGCGTCTGCGATGAGCG 1120  
 361 GlnArgLeuLeuAlaLeuValProMetGlyIleProArgThrLeuMetArgThrThrArg 380  
 1121 CAGCGGCTGTGGCGTGGTGGCCATCGGAATACCCGCGACCTCATTCGAGACACCCGCG 1180  
 381 PheArgGlyTyrThrLeuProGlnGlyThrGluValPheProLeuLeuGlySerIleLeu 400  
 1181 TTCGAGGTGTACACCTGCGCCAGGGCAGCGAGGTCTTCCCTCTCTCTGCTCTCATCTCTG 1240  
 401 HisAspProHisLysHisProGluGluPheAsnProAspArgPheLeuAspAla 420  
 1241 CATGACCCCAACATCTTCAAGACCCCAAGAGAGTTCAACCCAGACCGCTTCTCTGGATGCA 1300  
 421 AspGlyArgPheArgLysHisGluAlaPheLeuProPheSerLeuGlyLysArgValCys 440  
 1301 GATGGAGGTTCAGGAAGCATGAGGCGTTCCTGCGCTTCTCTTAGGGAAGCGTGTCTGTC 1360

QY 441 LeuGlyGluGlyLeuAlaLysAlaGluLeuPhePheThrThrLeuGlnAla 460  
 Db 1361 CTTGAGAGGGCTCGCAAGCGAGCTCTTCCTTCCACCACTCTCAAGCC 1420  
 QY 461 PheSerLeuGluSerProCysProProAspThrLeuSerLeuLysProThrValSerGly 480  
 Db 1421 TTCCTCCCTGGAGAGCCGCGCCGCGGACACCTCGAGCTCAAGCCACCGTCAGTGGC 1480  
 QY 481 LeuPheAenIleProProAlaPheGlnLeuGlnValArgProThrAspLeuHisSerThr 500  
 Db 1481 CTTTTCACACTTCCCCAGCCTTCCAGCTGCAGTCCGTCCTCCACTGCCTTCACTCCACC 1540  
 QY 501 ThrGlnThrArg 504  
 Db 1541 ACGCAGACCAGA 1552

RESULT 11  
 ABX89172  
 ID ABX89172 standard; cDNA; 2623 BP.  
 XX  
 AC ABX89172;  
 XX  
 DT 13-MAY-2003 (first entry)  
 XX  
 DE DNA encoding novel secreted and transmembrane protein PRO1906.  
 XX  
 KW Human; PRO; hypertrophy of neonatal heart; angiogenesis; wound healing;  
 KW cardiac insufficiency disorder; cancer; tumour; immune response;  
 KW adrenal cortical capillary endothelial growth; c-fos induction;  
 KW vascular endothelial growth factor inhibition; VEGF inhibition;  
 KW endothelial cell growth inhibitor; T-lymphocytes stimulation;  
 KW retinal neurons cell survival; rod photoreceptor cell survival;  
 KW retinal disorder; retinitis pigmentosa; kidney disorder;  
 KW mammalian kidney mesangial cell proliferation; Berger disease;  
 KW dermatitis; herpeticiformis; Crohn's disease; chondrocyte proliferation;  
 KW chondrocyte redifferentiation; sports injury; arthritis; gene; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US2003017563-A1.  
 XX  
 PD 23-JAN-2003.  
 XX  
 PF 07-MAY-2002; 2002US-00140808.  
 XX  
 PR 31-MAR-1997; 97WO-US005230.  
 PR 12-JUN-1998; 98WO-US012456.  
 PR 28-AUG-1998; 98WO-US014552.  
 PR 28-AUG-1998; 98WO-US017888.  
 PR 10-SEP-1998; 98WO-US018824.  
 PR 14-SEP-1998; 98WO-US019093.  
 PR 14-SEP-1998; 98WO-US019094.  
 PR 14-SEP-1998; 98WO-US019177.  
 PR 16-SEP-1998; 98WO-US019330.  
 PR 17-SEP-1998; 98WO-US019437.  
 PR 07-OCT-1998; 98WO-US021141.  
 PR 29-OCT-1998; 98WO-US022991.  
 PR 29-OCT-1998; 98WO-US022992.  
 PR 20-NOV-1998; 98WO-US024855.  
 PR 01-DEC-1998; 98WO-US025108.  
 PR 05-JAN-1999; 98WO-US000106.  
 PR 08-MAR-1999; 98WO-US005028.  
 PR 10-MAR-1999; 98WO-US005190.  
 PR 20-APR-1999; 98WO-US008615.  
 PR 14-MAY-1999; 98WO-US010733.  
 PR 02-JUN-1999; 98WO-US012252.  
 PR 01-SEP-1999; 98WO-US020111.  
 PR 08-SEP-1999; 98WO-US020594.  
 PR 13-SEP-1999; 98WO-US020944.  
 PR 15-SEP-1999; 98WO-US021090.  
 PR 15-SEP-1999; 98WO-US021547.  
 PR 05-OCT-1999; 98WO-US023089.  
 PR 29-NOV-1999; 98WO-US028214.  
 PR 30-NOV-1999; 99WO-US028313.  
 PR 30-NOV-1999; 99WO-US028409.  
 PR 01-DEC-1999; 99WO-US028301.  
 PR 01-DEC-1999; 99WO-US028634.  
 PR 02-DEC-1999; 99WO-US028551.  
 PR 02-DEC-1999; 99WO-US028584.  
 PR 02-DEC-1999; 99WO-US028585.  
 PR 16-DEC-1999; 99WO-US030095.  
 PR 20-DEC-1999; 99WO-US030911.  
 PR 20-DEC-1999; 99WO-US030999.  
 PR 22-DEC-1999; 99WO-US030720.  
 PR 30-DEC-1999; 99WO-US031243.  
 PR 30-DEC-1999; 99WO-US031274.  
 PR 05-JAN-2000; 2000WO-US000219.  
 PR 06-JAN-2000; 2000WO-US000277.  
 PR 06-JAN-2000; 2000WO-US000376.  
 PR 11-FEB-2000; 2000WO-US003565.  
 PR 18-FEB-2000; 2000WO-US004341.  
 PR 18-FEB-2000; 2000WO-US004342.  
 PR 22-FEB-2000; 2000WO-US004914.  
 PR 24-FEB-2000; 2000WO-US005004.  
 PR 01-MAR-2000; 2000WO-US005601.  
 PR 02-MAR-2000; 2000WO-US005746.  
 PR 02-MAR-2000; 2000WO-US005841.  
 PR 10-MAR-2000; 2000WO-US006319.  
 PR 15-MAR-2000; 2000WO-US006884.  
 PR 20-MAR-2000; 2000WO-US007377.  
 PR 21-MAR-2000; 2000WO-US007532.  
 PR 30-MAR-2000; 2000WO-US008439.  
 PR 17-MAY-2000; 2000WO-US013705.  
 PR 22-MAY-2000; 2000WO-US014042.  
 PR 30-MAY-2000; 2000WO-US014941.  
 PR 02-JUN-2000; 2000WO-US015264.  
 PR 28-JUL-2000; 2000WO-US020710.  
 PR 11-AUG-2000; 2000WO-US022031.  
 PR 23-AUG-2000; 2000WO-US023522.  
 PR 24-AUG-2000; 2000WO-US023328.  
 PR 08-NOV-2000; 2000WO-US030952.  
 PR 10-NOV-2000; 2000WO-US030873.  
 PR 01-DEC-2000; 2000WO-US032678.  
 PR 20-DEC-2000; 2000WO-US034956.  
 PR 28-FEB-2001; 2001WO-US0079498.  
 PR 01-MAR-2001; 2001WO-US006666.  
 PR 09-MAR-2001; 2001US-00802706.  
 PR 14-MAR-2001; 2001US-00808689.  
 PR 22-MAR-2001; 2001US-00816744.  
 PR 05-APR-2001; 2001US-00828366.  
 PR 10-MAY-2001; 2001US-00854208.  
 PR 18-MAY-2001; 2001US-00854280.  
 PR 25-MAY-2001; 2001US-00860216.  
 PR 25-MAY-2001; 2001US-00866028.  
 PR 25-MAY-2001; 2001US-00866034.  
 PR 01-JUN-2001; 2001WO-US017092.  
 PR 01-JUN-2001; 2001US-00872035.  
 PR 05-JUN-2001; 2001WO-US017800.  
 PR 14-JUN-2001; 2001US-00874503.  
 PR 19-JUN-2001; 2001US-00882636.  
 PR 20-JUN-2001; 2001US-00886342.  
 PR 21-JUN-2001; 2001WO-US019692.  
 PR 22-JUN-2001; 2001US-00887879.  
 PR 29-JUN-2001; 2001WO-US020116.  
 PR 09-JUL-2001; 2001WO-US021066.  
 PR 18-JUL-2001; 2001US-00908827.  
 PR 06-AUG-2001; 2001US-00924419.  
 PR 09-AUG-2001; 2001US-00927796.  
 PR 16-AUG-2001; 2001US-00931836.  
 PR 19-DEC-2001; 2001US-00028072.  
 XX  
 PA (GETH ) GENENTECH INC.





```

Db      1361  CTTGGAGAGGCGCTGCAAAAGCGAGGCTCTTCTCTTCAACCACTCCTCAAGCC 1420
Qv      461  PheSerLeuGluSerProCysProAspThrLeuSerLeuLysProThrValSerGly 480
          |||||
Db      1421  TTTCTCCCTGGAGAGCCGTCGCCCGGACACCCCTGAGCGCTCAAGCCACCGTCAGTGGC 1480
Qv      481  LeuPheAsnIleProProAlaPheGlnLeuGlnValArgProThrAspLeuHisSerThr 500
          |||||
Db      1481  CTTTTCACATTCGCCAGCCTTCCAGCTGCCAGTCCGTCCTCCACCTGACCTTCACCTCCACC 1540
Qv      501  ThrGlnThrArg 504
          |||||
Db      1541  ACGCAGACCAGA 1552

RESULT 12
ACD41826
ID      ACD41826 standard; cDNA; 2623 BP.
XX
AC      ACD41826;
XX
DT      05-SEP-2003 (first entry)
XX
DE      Human secreted/transmembrane protein (PRO) cDNA #32.
XX
KW      Human; ss; gene; PRO; secreted protein; transmembrane protein; tumour;
KW      cytostatic; gene therapy; tumour necrosis factor-alpha; TNF-alpha; blood;
KW      proteoglycan; cartilage; cytokine; peripheral blood mononuclear cell;
KW      PMBC; glucose uptake; FFA; skeletal muscle cell; adipocyte cell;
KW      chondrocyte cell proliferation; chondrocyte cell differentiation;
KW      pericyte cell; inner ear utricular supporting cell; T-lymphocyte cell;
KW      endothelial cell; A-peptide; factor VIIa.
XX
OS      Homo sapiens.
XX
PN      US2003036179 A1.
XX
PD      20-FEB-2003.
XX
PF      10-MAY-2002; 2002US-00142431.
XX
PR      31-MAR-1997; 97WO-US005230.
PR      12-JUN-1998; 98WO-US012456.
PR      14-JUL-1998; 98WO-US014552.
PR      28-AUG-1998; 98WO-US017888.
PR      10-SEP-1998; 98WO-US018824.
PR      14-SEP-1998; 98WO-US019093.
PR      14-SEP-1998; 98WO-US019094.
PR      16-SEP-1998; 98WO-US019177.
PR      17-SEP-1998; 98WO-US019330.
PR      07-OCT-1998; 98WO-US021141.
PR      29-OCT-1998; 98WO-US022991.
PR      20-NOV-1998; 98WO-US022992.
PR      01-DEC-1998; 98WO-US024855.
PR      05-JAN-1999; 99WO-US000106.
PR      08-MAR-1999; 99WO-US005028.
PR      10-MAR-1999; 99WO-US005190.
PR      20-APR-1999; 99WO-US008615.
PR      14-MAY-1999; 99WO-US010733.
PR      02-JUN-1999; 99WO-US012252.
PR      01-SEP-1999; 99WO-US020111.
PR      08-SEP-1999; 99WO-US020594.
PR      13-SEP-1999; 99WO-US020944.
PR      15-SEP-1999; 99WO-US021090.
PR      15-SEP-1999; 99WO-US021547.
PR      05-OCT-1999; 99WO-US023089.
PR      29-NOV-1999; 99WO-US028214.
PR      30-NOV-1999; 99WO-US028313.
PR      30-NOV-1999; 99WO-US028409.
PR      01-DEC-1999; 99WO-US028301.
PR      01-DEC-1999; 99WO-US028634.
PR      02-DEC-1999; 99WO-US028551.
PR      02-DEC-1999; 99WO-US028564.
PR      02-DEC-1999; 99WO-US028565.
PR      16-DEC-1999; 99WO-US030095.
PR      20-DEC-1999; 99WO-US030911.
PR      20-DEC-1999; 99WO-US030999.
PR      22-DEC-1999; 99WO-US030720.
PR      30-DEC-1999; 99WO-US031243.
PR      30-DEC-1999; 99WO-US031274.
PR      05-JAN-2000; 2000WO-US000219.
PR      06-JAN-2000; 2000WO-US000277.
PR      11-FEB-2000; 2000WO-US00376.
PR      11-FEB-2000; 2000WO-US003565.
PR      18-FEB-2000; 2000WO-US004341.
PR      22-FEB-2000; 2000WO-US004414.
PR      24-FEB-2000; 2000WO-US004914.
PR      24-FEB-2000; 2000WO-US005004.
PR      01-MAR-2000; 2000WO-US005601.
PR      02-MAR-2000; 2000WO-US005746.
PR      02-MAR-2000; 2000WO-US005841.
PR      10-MAR-2000; 2000WO-US006319.
PR      15-MAR-2000; 2000WO-US006884.
PR      20-MAR-2000; 2000WO-US007377.
PR      21-MAR-2000; 2000WO-US007532.
PR      30-MAR-2000; 2000WO-US008439.
PR      17-MAY-2000; 2000WO-US013705.
PR      22-MAY-2000; 2000WO-US014042.
PR      30-MAY-2000; 2000WO-US014941.
PR      02-JUN-2000; 2000WO-US015264.
PR      28-JUL-2000; 2000WO-US020710.
PR      11-AUG-2000; 2000WO-US022031.
PR      23-AUG-2000; 2000WO-US023522.
PR      24-AUG-2000; 2000WO-US023328.
PR      08-NOV-2000; 2000WO-US030952.
PR      10-NOV-2000; 2000WO-US030873.
PR      01-DEC-2000; 2000WO-US032678.
PR      20-DEC-2000; 2000US-00747259.
PR      20-DEC-2000; 2000WO-US034956.
PR      28-FEB-2001; 2001US-00796498.
PR      01-MAR-2001; 2001WO-US006520.
PR      09-MAR-2001; 2001WO-US006666.
PR      14-MAR-2001; 2001US-00802706.
PR      22-MAR-2001; 2001US-00808689.
PR      05-APR-2001; 2001US-00828366.
PR      10-MAY-2001; 2001US-00854208.
PR      18-MAY-2001; 2001US-00854280.
PR      25-MAY-2001; 2001US-00860216.
PR      25-MAY-2001; 2001US-00866026.
PR      25-MAY-2001; 2001US-00866034.
PR      01-JUN-2001; 2001WO-US017092.
PR      01-JUN-2001; 2001US-00872035.
PR      05-JUN-2001; 2001WO-US017800.
PR      14-JUN-2001; 2001US-00874503.
PR      19-JUN-2001; 2001US-00882636.
PR      20-JUN-2001; 2001US-00886342.
PR      21-JUN-2001; 2001WO-US019692.
PR      22-JUN-2001; 2001US-00887879.
PR      29-JUN-2001; 2001WO-US020116.
PR      09-JUL-2001; 2001WO-US021066.
PR      18-JUL-2001; 2001US-00908827.
PR      06-AUG-2001; 2001US-00924419.
PR      09-AUG-2001; 2001US-00927796.
PR      16-AUG-2001; 2001US-00931836.
PR      19-DEC-2001; 2001US-00028072.
XX
(GETH ) GENENTECH INC.
PA
XX
PI      Baker KP, Beresini M, Deforge L, Desnoyers L, Filvaroff E, Gao W;
PI      Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;
PI      Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;
XX

```



Db 1301 GATGAGCGTTGAGGAGCATGAGGCGTTCCTGCGCCCTTCTCTAGGAGCGTGTCTGC 1360  
 QY 441 LeuGlyGluGlyLeuAlaLysAlaGluLeuPhePheThrThrileLeuGlnAla 460  
 Db 1361 CTTGAGAGGGCTTGGCAAGGAGCTTCTCTCTTCCACCACTCCTCAAGCC 1420  
 QY 461 PheSerLeuGluSerProCysProProAspThrLeuSerLeuLysProThrValSerGly 480  
 Db 1421 TTCTCCTGGAGAGCGCTGCCCGCGGACACCTTGAGCCTCAAGCCACCGTCAAGTGC 1480  
 QY 481 LeuPheAsnIleProProAlaPheGlnLeuGlnValArgProThrAspLeuHisSerThr 500  
 Db 1481 CTTTTCACATTTCCCGAGCCTTCCAGCTGCAAGTCCGTCCCACTGACCTTCACTCCACC 1540  
 QY 501 ThrGlnThrArg 504  
 Db 1541 ACGCAGACCAGA 1552

RESULT 13  
 ACRA04055  
 ID ACA04055 standard; cDNA; 2623 BP.  
 XX  
 AC ACA04055;  
 XX  
 DT 27-MAY-2003 (first entry)  
 XX  
 DE Human cDNA encoding a secreted/transmembrane protein, SEQ ID 63.  
 XX  
 KW Human; ss; gene; secreted protein; transmembrane protein, PRO;  
 KW inflammatory disease; organ failure; atherosclerosis; cardiac injury;  
 KW infertility; birth defects; premature aging; AIDS; biosensor;  
 KW acquired immunodeficiency syndrome; cancer; diabetic complication;  
 KW bioreactor; tumour.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US2003032155-A1.  
 XX  
 PD 13-FEB-2003.  
 XX  
 PF 03-MAY-2002; 2002US-00137865.  
 XX  
 PR 31-MAR-1997; 98WO-US005230.  
 PR 12-JUN-1998; 98WO-US012456.  
 PR 14-JUL-1998; 98WO-US014552.  
 PR 28-AUG-1998; 98WO-US017888.  
 PR 10-SEP-1998; 98WO-US018824.  
 PR 14-SEP-1998; 98WO-US019093.  
 PR 14-SEP-1998; 98WO-US019094.  
 PR 16-SEP-1998; 98WO-US019330.  
 PR 17-SEP-1998; 98WO-US019437.  
 PR 07-OCT-1998; 98WO-US021141.  
 PR 29-OCT-1998; 98WO-US022991.  
 PR 29-OCT-1998; 98WO-US022992.  
 PR 20-NOV-1998; 98WO-US024855.  
 PR 01-DEC-1998; 98WO-US025108.  
 PR 05-JAN-1999; 98WO-US00106.  
 PR 08-MAR-1999; 98WO-US005028.  
 PR 10-MAR-1999; 98WO-US005190.  
 PR 20-APR-1999; 98WO-US008615.  
 PR 14-MAY-1999; 98WO-US010733.  
 PR 02-JUN-1999; 98WO-US012252.  
 PR 01-SEP-1999; 98WO-US020111.  
 PR 08-SEP-1999; 98WO-US020594.  
 PR 13-SEP-1999; 98WO-US020944.  
 PR 15-SEP-1999; 98WO-US021090.  
 PR 05-SEP-1999; 98WO-US021547.  
 PR 29-NOV-1999; 98WO-US023089.  
 PR 30-NOV-1999; 98WO-US028214.  
 PR 30-NOV-1999; 98WO-US028313.  
 PR 01-DEC-1999; 98WO-US028409.  
 PR 01-DEC-1999; 98WO-US028301.  
 PR 01-DEC-1999; 98WO-US028634.  
 PR 02-DEC-1999; 98WO-US028551.  
 PR 02-DEC-1999; 98WO-US028564.  
 PR 02-DEC-1999; 98WO-US028565.  
 PR 16-DEC-1999; 98WO-US030095.  
 PR 20-DEC-1999; 98WO-US030911.  
 PR 20-DEC-1999; 98WO-US030999.  
 PR 22-DEC-1999; 98WO-US030720.  
 PR 30-DEC-1999; 98WO-US031243.  
 PR 05-JAN-2000; 98WO-US031274.  
 PR 06-JAN-2000; 2000WO-US000219.  
 PR 06-JAN-2000; 2000WO-US000277.  
 PR 11-FEB-2000; 2000WO-US000376.  
 PR 18-FEB-2000; 2000WO-US004341.  
 PR 18-FEB-2000; 2000WO-US004342.  
 PR 22-FEB-2000; 2000WO-US004414.  
 PR 24-FEB-2000; 2000WO-US004914.  
 PR 24-FEB-2000; 2000WO-US005004.  
 PR 01-MAR-2000; 2000WO-US005601.  
 PR 02-MAR-2000; 2000WO-US005746.  
 PR 10-MAR-2000; 2000WO-US005841.  
 PR 15-MAR-2000; 2000WO-US006319.  
 PR 20-MAR-2000; 2000WO-US006884.  
 PR 21-MAR-2000; 2000WO-US007377.  
 PR 30-MAR-2000; 2000WO-US007532.  
 PR 17-MAY-2000; 2000WO-US008439.  
 PR 22-MAY-2000; 2000WO-US013705.  
 PR 30-MAY-2000; 2000WO-US014042.  
 PR 02-JUN-2000; 2000WO-US014941.  
 PR 28-JUL-2000; 2000WO-US015264.  
 PR 11-AUG-2000; 2000WO-US020710.  
 PR 23-AUG-2000; 2000WO-US023522.  
 PR 24-AUG-2000; 2000WO-US023328.  
 PR 08-NOV-2000; 2000WO-US030952.  
 PR 10-NOV-2000; 2000WO-US030873.  
 PR 01-DEC-2000; 2000WO-US032678.  
 PR 20-DEC-2000; 2000US-00747259.  
 PR 28-FEB-2001; 2001US-00796498.  
 PR 01-MAR-2001; 2001US-US006520.  
 PR 09-MAR-2001; 2001US-US006666.  
 PR 14-MAR-2001; 2001US-US0082706.  
 PR 22-MAR-2001; 2001US-US008689.  
 PR 05-APR-2001; 2001US-US016744.  
 PR 10-MAY-2001; 2001US-US028366.  
 PR 10-MAY-2001; 2001US-US0854208.  
 PR 18-MAY-2001; 2001US-US0854280.  
 PR 25-MAY-2001; 2001US-US0860216.  
 PR 25-MAY-2001; 2001US-US0866028.  
 PR 25-MAY-2001; 2001US-US0866034.  
 PR 01-JUN-2001; 2001US-US017092.  
 PR 01-JUN-2001; 2001US-US0872035.  
 PR 05-JUN-2001; 2001US-US017800.  
 PR 14-JUN-2001; 2001US-US0874503.  
 PR 19-JUN-2001; 2001US-US0882636.  
 PR 20-JUN-2001; 2001US-US0886342.  
 PR 21-JUN-2001; 2001US-US019692.  
 PR 22-JUN-2001; 2001US-US0887879.  
 PR 22-JUN-2001; 2001US-US0202116.  
 PR 09-JUL-2001; 2001US-US021066.  
 PR 18-JUL-2001; 2001US-US021735.  
 PR 06-AUG-2001; 2001US-US0908827.  
 PR 09-AUG-2001; 2001US-US0924419.  
 PR 16-AUG-2001; 2001US-US0927796.  
 PR 19-DEC-2001; 2001US-US0931836.  
 PR 19-DEC-2001; 2001US-US0028072.  
 XX  
 PA (GETH ) GENENTECH INC.  
 XX  
 PI Baker KP, Beresini M, Deforge L, Desnoyers L, Filvaroff E, Gao W;  
 PI Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;

Smith V, Stewart TA, Tunas D, Watanabe CK, Wood WI, Zhang Z;  
WPI; 2003-331925/31.  
P-PSDB; ABU66877.  
New secreted and transmembrane nucleic acids and polypeptides, designated  
as PRO, useful for treating inflammation, organ failure, atherosclerosis,  
cardiac injury, infertility, birth defects, premature aging, AIDS, or  
cancer.  
Claim 2; Fig 63; 659pp; English.  
The invention relates to an isolated nucleic acid comprising, or which is  
at least 80% identical to, or the full-length coding sequence of, any of  
the 275 nucleotide sequences, encoding the corresponding PRO polypeptide  
(one of 275 secreted or transmembrane proteins). The nucleic acid further  
comprises the full-length coding sequence of the DNA deposited under  
American Type Culture Collection (ATCC) accession number in a list given  
in the specification. Also included are vectors and host cells for  
producing PRO proteins, PRO fusion proteins, anti-PRO antibodies, PRO  
extracellular domains and mature sequences, methods of detecting PRO  
proteins, methods for stimulating the release of TNF-alpha (tumour  
necrosis factor alpha) from human blood, (and the proliferation of, or gene  
differentiation of chondrocyte cells, the release or proteoglycans from  
expression in pericyte cells, the release or proteoglycans from  
cartilage, proliferation of inner ear utricular supporting cells, the  
proliferation of T-lymphocyte cells, the release of a cytokine from  
peripheral blood mononuclear cells (PBMC), or the proliferation of  
endothelial cells, a method for modulating the uptake of glucose or free  
fatty acid (FFA) by skeletal muscle cells, a method for inhibiting the  
binding of A-peptide to factor VIIa, or the differentiation of adipocyte  
cells, a method for detecting the presence of a tumour in a mammal and an  
oligonucleotide probe derived from any of the nucleotide sequences cited  
above. The nucleic acids and polypeptides are useful for treating  
inflammatory diseases, organ failure, atherosclerosis, cardiac injury,  
infertility, birth defects, premature aging, AIDS (acquired  
immunodeficiency syndrome), cancer, or diabetic complications. The  
nucleic acids are useful as hybridisation probes, in chromosome and gene  
mapping, and in generating antisense RNA or DNA. The polypeptides are  
useful as pharmaceuticals, diagnostics, biosensors or bioreactors. Both  
are useful in tissue typing. The present sequence encodes a PRO protein  
of the invention

Sequence 2623 BP; 552 A; 865 C; 704 G; 502 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 5,37e-231 Length: 2623  
Score: 2615.00 Matches: 504  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 7 Gaps: 0

US-10-669-693-2 (1-504) x ACA04055 (1-2623)

QY 1 MetGluAlaThrGlyThrTroAlaLeuLeuLeuAlaLeuAlaLeuLeuLeuLeuLeuThr 20  
Db 41 ATGAGAGCGACCGGCGACCTGGGCGCTGCTGTGGCGCTGGCGCTGCTCTCTGCTGACG 101  
QY 21 LeuAlaLeuSerGlyThrArgAlaArgGlyHisLeuProProGlyProThrProLeuPro 40  
Db 101 CTGGCGCTGTCCGGACCGAGCCCGAGCCCGAGCCCGAGCCCGAGCCCGAGCCCGAGCC 161  
QY 41 LeuLeuGlyAsnLeuLeuGlnLeuArgProGlyAlaLeuTyrSerGlyLeuMetArgLeu 60  
Db 161 CTGCTGGGAAACCTCTCTGCAGCTACGGCCCGGGCGCTGTATTAGGCGCTCATGGCGGTG 221  
QY 61 SerIysIysTyrGlyProValPheThrIleTyrLeuGlyProTyrArgProValVal 80  
Db 221 AGTAAGAAGTACGACCGGTGTTACCATCTACCTGGAGCCCTGGCGGCTGTGTGTGTC 281  
QY 81 LeuValGlyGlnGluAlaValArgGluAlaLeuGlyGlyGlnAlaGluGluPheSerGly 101

```

Oy 461 PheSerLeuGluSerProCysProAspThrLeuSerLeuValSerProThrValSerGly 480
Db 1421 TTCTCCCTGGAGAGCCCTGCCCGGACACCCCTGAGCCTCAAGCCACCGTCAGTGGC 1480
Oy 481 LeuPheAsnIleProAlaPheGlnLeuGlnValArgProThrAspLeuHisSerThr 500
Db 1481 CTTTTCACATTCCCGCAGCCTTCCAGCTGCAAGTCCGTCCTCCACTGACCTTCACCTCCACC 1540
Oy 501 ThrGlnThrArg 504
Db 1541 ACGCAGACCAGA 1552
RESULT 14
ADA45582
ID ADA45582 standard; cDNA; 2623 BP.
XX
AC ADA45582;
XX
DT 20-NOV-2003 (first entry)
XX
DE Novel human secreted and transmembrane protein PRO1906 cDNA.
XX
KW Human; secreted and transmembrane protein; PRO; gene; ss;
KW Tumour necrosis factor alpha release; TNF-alpha release;
KW glucose uptake modulator; FFA uptake modulator;
KW cell proliferation stimulator; cell differentiation stimulator;
KW cell differentiation inhibitor; cytokine release stimulator; tumour;
KW lung tumour; colon tumour; breast tumour; prostate tumour; rectal tumour;
KW cervical tumour; liver tumour; chromosome mapping; gene mapping;
KW gene therapy; chromosome identification; chromosome marker.
XX
OS Homo sapiens.
XX
PN US2003022328-A1.
XX
PD 30-JAN-2003.
XX
PF 16-APR-2002; 2002US-00123904.
XX
PR 31-MAR-1997; 99WO-US005230.
PR 12-JUN-1998; 98WO-US012456.
PR 14-JUL-1998; 98WO-US014552.
PR 28-AUG-1998; 98WO-US017888.
PR 10-SEP-1998; 98WO-US018824.
PR 14-SEP-1998; 98WO-US019093.
PR 14-SEP-1998; 98WO-US019094.
PR 14-SEP-1998; 98WO-US019177.
PR 16-SEP-1998; 98WO-US019330.
PR 17-SEP-1998; 98WO-US019437.
PR 07-OCT-1998; 98WO-US021141.
PR 29-OCT-1998; 98WO-US022991.
PR 29-OCT-1998; 98WO-US022992.
PR 01-DEC-1998; 98WO-US024855.
PR 05-JAN-1999; 98WO-US025108.
PR 08-MAR-1999; 98WO-US005028.
PR 10-MAR-1999; 98WO-US005190.
PR 20-APR-1999; 98WO-US008615.
PR 14-MAY-1999; 98WO-US010733.
PR 02-JUN-1999; 98WO-US012252.
PR 01-SEP-1999; 98WO-US020111.
PR 08-SEP-1999; 98WO-US020594.
PR 13-SEP-1999; 98WO-US020944.
PR 15-SEP-1999; 98WO-US021090.
PR 15-SEP-1999; 98WO-US021547.
PR 05-OCT-1999; 98WO-US023089.
PR 29-NOV-1999; 98WO-US028214.
PR 30-NOV-1999; 98WO-US028313.
PR 01-DEC-1999; 98WO-US028409.
PR 01-DEC-1999; 98WO-US028301.
PR 02-DEC-1999; 98WO-US028634.
PR 02-DEC-1999; 98WO-US028551.
PR 02-DEC-1999; 99WO-US028564.
PR 02-DEC-1999; 99WO-US028565.
PR 16-DEC-1999; 99WO-US030095.
PR 20-DEC-1999; 99WO-US030911.
PR 20-DEC-1999; 99WO-US030999.
PR 22-DEC-1999; 99WO-US030999.
PR 30-DEC-1999; 99WO-US030720.
PR 30-DEC-1999; 99WO-US031243.
PR 05-JAN-2000; 99WO-US031274.
PR 05-JAN-2000; 2000WO-US000219.
PR 06-JAN-2000; 2000WO-US000277.
PR 11-FEB-2000; 2000WO-US000376.
PR 18-FEB-2000; 2000WO-US003565.
PR 18-FEB-2000; 2000WO-US004341.
PR 22-FEB-2000; 2000WO-US004342.
PR 24-FEB-2000; 2000WO-US004414.
PR 24-FEB-2000; 2000WO-US004914.
PR 01-MAR-2000; 2000WO-US005004.
PR 02-MAR-2000; 2000WO-US005601.
PR 02-MAR-2000; 2000WO-US005746.
PR 10-MAR-2000; 2000WO-US005841.
PR 15-MAR-2000; 2000WO-US006319.
PR 20-MAR-2000; 2000WO-US006884.
PR 21-MAR-2000; 2000WO-US007377.
PR 30-MAR-2000; 2000WO-US007532.
PR 17-MAY-2000; 2000WO-US013705.
PR 30-MAY-2000; 2000WO-US014042.
PR 02-JUN-2000; 2000WO-US014941.
PR 28-JUL-2000; 2000WO-US015264.
PR 11-AUG-2000; 2000WO-US020710.
PR 23-AUG-2000; 2000WO-US022031.
PR 24-AUG-2000; 2000WO-US023522.
PR 08-NOV-2000; 2000WO-US030952.
PR 01-DEC-2000; 2000WO-US030873.
PR 20-DEC-2000; 2000WO-US032578.
PR 28-FEB-2001; 2000WO-US034956.
PR 01-MAR-2001; 2001WO-US006520.
PR 09-MAR-2001; 2001WO-US006666.
PR 14-MAR-2001; 2001US-00802706.
PR 22-MAR-2001; 2001US-00816744.
PR 05-APR-2001; 2001US-00828366.
PR 10-MAY-2001; 2001US-00854208.
PR 10-MAY-2001; 2001US-00854280.
PR 18-MAY-2001; 2001US-00860216.
PR 25-MAY-2001; 2001US-00866028.
PR 25-MAY-2001; 2001US-00866034.
PR 01-JUN-2001; 2001US-00872035.
PR 01-JUN-2001; 2001WO-US017800.
PR 05-JUN-2001; 2001US-00874503.
PR 14-JUN-2001; 2001US-00882636.
PR 19-JUN-2001; 2001US-00886342.
PR 20-JUN-2001; 2001WO-US019692.
PR 21-JUN-2001; 2001US-00887879.
PR 22-JUN-2001; 2001WO-US020116.
PR 29-JUN-2001; 2001WO-US021066.
PR 09-JUL-2001; 2001WO-US021735.
PR 18-JUL-2001; 2001US-00908827.
PR 06-AUG-2001; 2001US-00924419.
PR 09-AUG-2001; 2001US-00927796.
PR 16-AUG-2001; 2001US-00931836.
PR 19-DEC-2001; 2001US-00028072.
XX
XX
PA (GETH ) GENENTECH INC.
XX
PI Baker KP, Beresini M, Deforge L, Desnoyers L, Filvaroff E, Gao W;
PI Gerritsen ME, Goddard A, Godowski PU, Gurney AL, Sherwood S;
PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;
XX

```

DR WPI; 2003-584997/55.  
 DR P-PSDB; ADA45583.  
 XX  
 PT Novel secreted and transmembrane polypeptide for modulating biological  
 PT activity of cell expressing the polypeptide, identifying agonists or  
 PT antagonists of polypeptide, and as molecular weight markers.  
 XX  
 XX Claim 2; Fig 63; 659pp; English.  
 XX  
 CC The invention describes 305 nucleic acids encoding PRO (secreted and  
 CC transmembrane) polypeptides (I). (I) is useful for stimulating the  
 CC release of TNF-alpha from human blood, for modulating the uptake of  
 CC glucose or FFA by skeletal muscle cells or adipocyte cells, for  
 CC stimulating the proliferation or differentiation of chondrocyte cells,  
 CC for stimulating the proliferation of or gene expression in pericyte  
 CC cells, for stimulating the release of proteoglycans from cartilage, for  
 CC stimulating the proliferation of inner ear utricular supporting cells,  
 CC for stimulating the proliferation of T-lymphocyte cells, for stimulating  
 CC the release of a cytokine from PBMC cells, for inhibiting the binding of  
 CC A-peptide to factor VIIa, for inhibiting the differentiation of adipocyte  
 CC cells, for stimulating proliferation of endothelial cells, for detecting  
 CC the presence of tumour in a mammal. The tumour is lung, colon, breast,  
 CC prostate, rectal, cervical or liver tumour. The oligonucleotide probes  
 CC are useful for isolating genomic and cDNA nucleotide sequences or  
 CC antisense probes. (I) is also useful as therapeutic agent. PRO is useful  
 CC in assays to identify other proteins or molecules involved in binding  
 CC interaction. A polynucleotide (II) encoding (I) is useful in chromosome  
 CC and gene mapping, in generation of antisense RNA and DNA, in the  
 CC preparation of PRO polypeptide, for generating transgenic animals or  
 CC knockout animals which in turn are useful in the development and  
 CC screening of therapeutically useful reagents, in gene therapy, for  
 CC chromosome identification, as chromosome marker, and for generating  
 CC probes. An anti-(I)-antibody is useful in diagnostic assays for PRO, e.g.  
 CC detecting its expression in specific cells, tissues or serum, and for  
 CC affinity purification of PRO from recombinant cell culture or natural  
 CC sources. (I) and (II) are useful for tissue typing. This sequence encodes  
 CC a novel human secreted and transmembrane PRO polypeptide.  
 XX  
 SQ Sequence 2623 BP; 552 A; 865 C; 704 G; 502 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 5,37e-231 Length: 2623  
 Score: 2615.00 Matches: 504  
 Percent Similarity: 100.00% Conservatives: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: Gaps: 0  
 US-10-669-693-2 (1-504) x ADA45582 (1-2623)  
 QY 1 MetGluAlaThrGlyThrTrpAlaLeuLeuLeuAlaLeuLeuLeuLeuLeuThr 20  
 DB 41 ATGGAGCGCACCGGCACCTGGCGCTGCTGGCGCTGGCGCTGCTGGCGCTGACG 100  
 QY 21 LeuAlaLeuSerGlyThrArgAlaArgGlyHisLeuProProGlyProThrProLeuPro 40  
 DB 101 CTGGCGCTGTCGGGACACGAGGCGCCGAGGCCACTGCCCCCGGGCCACCGCGCTACCA 160  
 QY 41 LeuLeuGlyAsnLeuLeuGlnLeuArgProGlyAlaLeuTyrSerGlyLeuMetArgLeu 60  
 DB 161 CTGCTGGAAACCTCCTGCAGCTACGCCCGGGGCGCTGTATTTCAGGGCTCATGGCGCTG 220  
 QY 61 SerLysIysTyrGlyProValPheThrIleTyrLeuGlyProThrArgProValValVal 80  
 DB 221 AGTAAGAAGTACGAGCGGTGTTCACCATCTACCTGGGACCCCTGGGCGCGCTGGTGGTC 280  
 QY 81 LeuValGlyGlnGluAlaValArgGluAlaLeuGlyGlyGlnAlaGluGluPheSerGly 100  
 DB 281 CTGGTGGGAGGAGGCTGTGGGAGGCGCTGGAGGTCAGGCTGAGGAGTTTCAGCGGC 340  
 QY 101 ArgGlyThrValAlaMetLeuGluGlyThrPheAspGlyHisGlyValPhePheSerAsn 120  
 DB 341 CGGGGAACCGTAGCGATGCTGGAAGGACTTTTGTATGCCATGGGGTTTCTCTCTCCAAC 400

121 GlyGluArgTrpArgGlnLeuArgLysPheThrMetLeuAlaLeuArgAspLeuGlyMet 140  
 DB 401 GGGAGACGGTGGAGGACGCTGAGGAGTTTACCATGCTTGTCTCGGGACCTGGGCATG 460  
 QY 141 GlyLysArgGluGlyGluLeuLeuGlnAlaGluAlaArgCysLeuValGluThrPhe 160  
 DB 461 GGGAGCCGAGAAGCGGAGAGCTGATCAGGCGGAGGCGCGGTCTCTGTGGAGCATTC 520  
 QY 161 GlnGlyThrGluGlyArgProPheAspProSerLeuLeuAlaGlnAlaThrSerAsn 180  
 DB 521 CAGGGGACAGAGAAGGACGCCCATTCGATCCTCCCTGCTGGCCAGGCCACCTCCCAAC 580  
 QY 181 ValValCysSerLeuLeuPheGlyLeuArgPheSerTyrGluAspLysGluPheGlnAla 200  
 DB 581 GTAGTCTGCTCCCTCTCTCTTGGCCCTCGCTTCTCTATAGGATAAGGAGTTCCAGGCC 640  
 QY 201 ValValArgAlaAlaGlyGlyThrLeuLeuGlyValSerSerGlnGlyGlnThrTyr 220  
 DB 641 GTGGTCCGGGACGCTGGTGGTACCTGCTGGAGTCAGCTCCAGGGGGGTCCAGACTAC 700  
 QY 221 GluMetPheSerTrpPheLeuArgProLeuProGlyProHisLysGlnLeuLeuHis 240  
 DB 701 GAGATGTTCTCTCTGGTTCCTGGGCCCTGCGGAGGCTCCAGGAGGCTCCCTCCACC 760  
 QY 241 ValSerThrLeuAlaAlaPheThrValArgGlnValGlnGlnHisGlnGlyAsnLeuAsp 260  
 DB 761 GTCAGCAGCTTGGCTGGCTTCCAGTCCGGCAGGTGCAGCAGCAGCAGGAGAACCTGGAT 820  
 QY 261 AlaSerGlyProAlaArgAspLeuValAspAlaPheLeuLeuLysMetAlaGlnGluGlu 280  
 DB 821 GCTTCGGGCCCGGCACGTCGATGCTTCTGCTGAAAGTGCACAGGAGGAA 880  
 QY 281 GlnAsnProGlyThrGluPheThrLysAsnMetLeuMetThrValIleTyrLeuLeu 300  
 DB 881 CAAACCCAGGACAGAAATTCACCAAGAAACATGCTGATGACAGTCAATTTATTGCTG 940  
 QY 301 PheAlaGlyThrMetThrValSerThrValGlyTyrThrLeuLeuLeuLeuMetLys 320  
 DB 941 TTTGCTGGAGCAGTACGCTGAGCAGCAGCAGCAGCTATACCTCTCTCTGATGATAA 1000  
 QY 321 TyrProHisValGlnLysTrpValArgGluGluLeuAsnArgGluLeuGlyAlaGlyGln 340  
 DB 1001 TACCCTCATGTCCAAAAGTGGGTACGTGAGGAGCTGAATCGGAGCTGGGGGTGGCCAG 1060  
 QY 341 AlaProSerLeuGlyAspArgThrArgLeuProTyrThrAspAlaValLeuHisGluAla 360  
 DB 1061 GCACCAAGCTTAGGGGACCGTACCCGCTTACACGACGCGGTTCGCAATGAGGCG 1120  
 QY 361 GlnArgLeuLeuAlaLeuValProMetGlyIleProArgThrLeuMetArgThrThrArg 380  
 DB 1121 CAGCGGCTGCTGGCGTGGTCCCATGGGAATACCCCGCAGCCTCATTCGGGACACCCGC 1180  
 QY 381 PheArgGlyTyrThrLeuProGlnGlyThrGluValPheProLeuLeuGlySerIleLeu 400  
 DB 1181 TCCGAGGCTTACACCTGCCCCCAGGAGGCTTCCCTCTCTCTGCTCCATCTCTG 1240  
 QY 401 HisAspProAsnIlePheLysHisProGluGluPheAsnProAspArgPheLeuAspAla 420  
 DB 1241 CATGACCCCAACATCTTCAGGACCCCAAGAGATTCAACCCAGACCGTTTCTTGGATGCA 1300  
 QY 421 AspGlyArgPheArgLysHisGluAlaPheLeuProPheSerLeuGlyLysArgValCys 440  
 DB 1301 GATGGACGGTTACGGAAGCATGAGGCGTTCCTGCCCTTCTCTTAGGGAAGCGTCTGC 1360  
 QY 441 LeuGlyGluGlyLeuAlaLysAlaGluLeuPhePhePheThrThrIleLeuGlnAla 460  
 DB 1361 CTTGGAGAGGCGCTGGCAAAAGCGAGCTCTTCTCTTCTTCCACCATCTTCCATCAGGCC 1420  
 QY 461 PheSerLeuGluSerProCysProProAspThrLeuSerLeuLysProThrValSerGly 480  
 DB 1421 TTTCTCTCTGGAGAGCCCGTGGCCCGGACACCTTGAGCCTCAAGCCACCGTCAGTGGC 1480



QY 481 LeuPheAsnIleProProAlaPheGlnLeuGlnValArgProThrAspLeuHisSerThr 500  
Db 1481 CTTTTCACATTCCTCCAGCCTTCCAGCTGCAAGTCCGTCCTCCACTGACCTTCCTCCACC 1540  
QY 501 ThrGlnThrArg 504  
Db 1541 ACGCAGACCAGA 1552  
RESULT 15  
ADA76013  
ID ADA76013 standard; cDNA; 2623 BP.  
XX  
AC ADA76013;  
XX  
DT 20-NOV-2003 (first entry)  
XX  
DE Human PRO polynucleotide #32.  
XX  
KW Human; gene; ss; PRO; secreted polypeptide; transmembrane polypeptide;  
KW tumour necrosis factor-alpha; TNF-alpha; chondrocyte cell; tumour;  
KW cancer; adrenal; lung; colon; breast; prostate; rectum; kidney; cervix;  
KW liver; microvascular endothelial cell; glucose; FFA;  
KW skeletal muscle cell; adipocyte cell; pericyte cell;  
KW inner ear utricular supporting cell; T-lymphocyte cell;  
KW endothelial cell tube formation; bone disorder; cartilage disorder;  
KW sports injury; proteoglycan; articular cartilage defect; osteoarthritis;  
KW rheumatoid arthritis; haemoglobin-associated disorder thalassaemia;  
KW immune system cell infiltration.  
XX  
OS Homo sapiens.  
XX  
PN US2003073212-A1.  
XX  
PD 17-APR-2003.  
XX  
PF 16-APR-2002; 2002US-00123303.  
XX  
PR 31-MAR-1997; 97WO-US005230.  
PR 12-JUN-1998; 98WO-US012456.  
PR 14-JUL-1998; 98WO-US014552.  
PR 28-AUG-1998; 98WO-US017888.  
PR 10-SEP-1998; 98WO-US018824.  
PR 14-SEP-1998; 98WO-US019093.  
PR 14-SEP-1998; 98WO-US019094.  
PR 14-SEP-1998; 98WO-US019177.  
PR 16-SEP-1998; 98WO-US019330.  
PR 17-SEP-1998; 98WO-US019437.  
PR 07-OCT-1998; 98WO-US021141.  
PR 29-OCT-1998; 98WO-US022991.  
PR 29-OCT-1998; 98WO-US022992.  
PR 01-NOV-1998; 98WO-US024855.  
PR 01-DEC-1998; 98WO-US025108.  
PR 05-JAN-1999; 98WO-US000106.  
PR 08-MAR-1999; 98WO-US005028.  
PR 10-MAR-1999; 98WO-US005190.  
PR 20-APR-1999; 98WO-US008615.  
PR 14-MAY-1999; 98WO-US010733.  
PR 02-JUN-1999; 98WO-US012252.  
PR 01-SEP-1999; 98WO-US020111.  
PR 08-SEP-1999; 98WO-US020594.  
PR 13-SEP-1999; 98WO-US020944.  
PR 15-SEP-1999; 98WO-US021090.  
PR 15-SEP-1999; 98WO-US021547.  
PR 05-OCT-1999; 98WO-US023089.  
PR 29-NOV-1999; 98WO-US028214.  
PR 30-NOV-1999; 98WO-US028313.  
PR 30-NOV-1999; 98WO-US028409.  
PR 01-DEC-1999; 98WO-US028301.  
PR 01-DEC-1999; 98WO-US028634.  
PR 02-DEC-1999; 98WO-US028551.  
PR 02-DEC-1999; 98WO-US028564.  
PR 02-DEC-1999; 98WO-US028565.  
PR 16-DEC-1999; 98WO-US030095.  
PR 20-DEC-1999; 99WO-US030911.  
PR 20-DEC-1999; 99WO-US030999.  
PR 22-DEC-1999; 99WO-US030720.  
PR 30-DEC-1999; 99WO-US031243.  
PR 30-DEC-1999; 99WO-US031274.  
PR 05-JAN-2000; 2000WO-US000219.  
PR 06-JAN-2000; 2000WO-US000277.  
PR 06-JAN-2000; 2000WO-US000376.  
PR 11-FEB-2000; 2000WO-US003565.  
PR 18-FEB-2000; 2000WO-US004341.  
PR 18-FEB-2000; 2000WO-US004342.  
PR 22-FEB-2000; 2000WO-US004414.  
PR 24-FEB-2000; 2000WO-US004914.  
PR 24-FEB-2000; 2000WO-US005004.  
PR 01-MAR-2000; 2000WO-US005601.  
PR 02-MAR-2000; 2000WO-US005746.  
PR 02-MAR-2000; 2000WO-US005841.  
PR 10-MAR-2000; 2000WO-US006319.  
PR 15-MAR-2000; 2000WO-US006884.  
PR 20-MAR-2000; 2000WO-US007377.  
PR 21-MAR-2000; 2000WO-US007532.  
PR 30-MAR-2000; 2000WO-US008439.  
PR 17-MAY-2000; 2000WO-US013705.  
PR 22-MAY-2000; 2000WO-US014042.  
PR 30-MAY-2000; 2000WO-US014941.  
PR 02-JUN-2000; 2000WO-US015264.  
PR 28-JUL-2000; 2000WO-US020710.  
PR 11-AUG-2000; 2000WO-US022031.  
PR 23-AUG-2000; 2000WO-US023522.  
PR 24-AUG-2000; 2000WO-US023328.  
PR 08-NOV-2000; 2000WO-US030952.  
PR 10-NOV-2000; 2000WO-US030873.  
PR 01-DEC-2000; 2000WO-US032678.  
PR 20-DEC-2000; 2000US-00747259.  
PR 20-DEC-2000; 2000WO-US034956.  
PR 28-FEB-2001; 2001US-00796498.  
PR 28-FEB-2001; 2001WO-US006520.  
PR 01-MAR-2001; 2001WO-US006666.  
PR 09-MAR-2001; 2001US-00802706.  
PR 14-MAR-2001; 2001US-00808689.  
PR 22-MAR-2001; 2001US-00816744.  
PR 05-APR-2001; 2001US-00828366.  
PR 10-MAY-2001; 2001US-00854208.  
PR 18-MAY-2001; 2001US-00860216.  
PR 25-MAY-2001; 2001US-00866028.  
PR 25-MAY-2001; 2001US-00866034.  
PR 25-MAY-2001; 2001WO-US017092.  
PR 01-JUN-2001; 2001US-00872035.  
PR 01-JUN-2001; 2001WO-US017800.  
PR 05-JUN-2001; 2001US-00874503.  
PR 14-JUN-2001; 2001US-00882636.  
PR 19-JUN-2001; 2001US-00886342.  
PR 21-JUN-2001; 2001WO-US019692.  
PR 21-JUN-2001; 2001US-00887879.  
PR 22-JUN-2001; 2001WO-US020116.  
PR 29-JUN-2001; 2001WO-US021066.  
PR 09-JUL-2001; 2001WO-US021735.  
PR 18-JUL-2001; 2001US-00908827.  
PR 06-AUG-2001; 2001US-00924419.  
PR 09-AUG-2001; 2001US-00927796.  
PR 16-AUG-2001; 2001US-00931836.  
PR 19-DEC-2001; 2001US-00028072.  
XX  
PA (GETH ) GENENTECH INC.  
XX  
PI Baker KP, Beresini M, Deforge L, Desnoyers L, Filvaroff E, Gao W;  
PI Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;  
PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;  
XX  
DR WPI; 2003-687639/65.  
DR P-PSDB; ADA76014.  
XX



Db 1421 TTCTCCTGGAGAGCCCGTGCCTGGGACACCTTGAGCTCAAGCCACCGTCAGTGGC 1480  
 QY 481 LeuPheAsnIleProProAlaPheGlnLeuGlnValArgProThrAspLeuHisSerThr 500  
 Db 1481 CTTTCAACATTCCCGCCAGCCCTCCAGCTGCAAGTCCGTCCTGACCTTCACTCCACC 1540  
 QY 501 ThrGlnThrArg 504  
 Db 1541 ACGCAGACCAGA 1552

Search completed: September 18, 2004, 01:51:19  
 Job time : 677 secs

***This Page Blank (uspto)***

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 18, 2004, 00:15:11 ; Search time 3778 Seconds  
(without alignments)  
3983.732 Million cell updates/sec

Title: US-10-669-693-2  
Perfect score: 2615  
Sequence: 1 MEATGTWALLALALLLLT.....PPAFQLQVRPTDLHSTTQTR 504

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO.spool/US10669693/runat\_15092004\_101310\_5862/app\_query.fasta\_1.647  
-DB=EST -QFMT=fastcap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USR=US10669693 @CGN 1 1 3437 @runat\_15092004\_101310\_5862 -NCPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estmu:\*  
4: em\_estnu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_htc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_est3:\*  
12: gb\_est4:\*  
13: gb\_est5:\*  
14: gb\_est6:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_man:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rod:\*  
26: em\_gss\_pig:\*  
27: em\_gss\_vrl:\*  
28: gb\_gsl1:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	2060.5	78.8	2292	11	AK087069	AK087069 Mus muscu
2	2060.5	78.8	2621	11	AK004699	AK004699 Mus muscu
3	1529.5	58.5	900	13	BUS52396	BUS52396 AGENCOURT
4	1465.5	56.0	924	13	B0849902	B0849902 AGENCOURT
5	1447.5	55.3	903	13	BQ900302	BQ900302 AGENCOURT
6	1410.5	53.9	939	13	BQ958112	BQ958112 AGENCOURT
7	1376.5	52.6	2742	11	AK054324	AK054324 Mus muscu
8	1362.5	52.1	857	12	BM047426	BM047426 603628712
9	1321.5	50.5	877	13	BUS43186	BUS43186 AGENCOURT
10	1299.5	49.7	1876	11	AK028103	AK028103 Mus muscu
11	1222.5	46.7	838	13	BQ957100	BQ957100 AGENCOURT
12	1146.5	43.8	1944	11	BC022141	BC022141 Mus muscu
13	1138.5	43.5	969	13	EX422464	EX422464 BX422464
14	1083.5	41.4	2870	11	AK050051	AK050051 Mus muscu
15	1069.5	40.9	1870	11	AK008688	AK008688 Mus muscu
16	1056.5	40.4	1825	11	AK005017	AK005017 Mus muscu
17	1039.5	39.8	1567	14	CD013962	CD013962 90137726
18	1038.5	39.7	1990	11	AK008580	AK008580 Mus muscu
19	1022.5	39.1	723	12	BI906340	BI906340 603063307
20	1022.5	39.1	2022	14	CD014119	CD014119 90127964
21	1018.5	38.9	2093	14	CD013967	CD013967 90127924
22	1007.5	38.5	1452	14	CD013965	CD013965 90137857
23	1005.5	38.4	2131	14	CD013968	CD013968 90127932
24	1001.5	38.3	2132	14	CD013969	CD013969 90127948
25	998.5	38.2	1559	11	AK050436	AK050436 Mus muscu
26	993.5	38.0	1482	29	AY407531	AY407531 Mus muscu
27	993.5	38.0	1731	11	BC042693	BC042693 Mus muscu
28	992.5	37.9	599	12	BM786289	BM786289 K-EST0064
29	982.5	37.6	1482	29	AY407529	AY407529 Homo sapi
30	972.5	37.2	1703	11	AK050082	AK050082 Mus muscu
31	954.5	36.5	1339	14	CD013963	CD013963 90137742
32	953.5	36.4	1323	14	CD013966	CD013966 90137742
33	943.5	36.1	1299	14	CD013964	CD013964 90137766
34	936.5	35.8	1760	14	CD013970	CD013970 90127846
35	933.5	35.7	571	10	BE148597	BE148597 MR0-HT024
36	928.5	35.5	648	12	BI016305	BI016305 MR4-ET013
37	918.5	35.1	674	12	BI836998	BI836998 603084539
38	911.5	34.9	669	13	BQ257072	BQ257072 NISC_K009
39	905.5	34.6	1773	11	AK008512	AK008512 Mus muscu
40	903.5	34.5	538	12	BM795225	BM795225 K-EST0076
41	893.5	34.1	548	13	BM646408	BM646408 DXFZP781K
42	887.5	33.9	637	14	CF170039	CF170039 B0821G03-
43	887.5	33.9	553	12	BM751056	BM751056 K-EST0026
44	870.5	33.3	626	14	CF170325	CF170325 B0825H09-
45	867.5	33.2	804	12	BM044172	BM044172 603621408

# ALIGNMENTS

## RESULT 1

AK087069

LOCUS

DEFINITION

AK087069

VERSION

KEYWORDS

SOURCE

ORGANISM

AK087069

AK087069.1

GI:26352314

HTC; CAP trapper.

Mus musculus (house mouse)

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

AK087069 2292 bp mRNA linear HTC 20-SEP-2003  
Mus musculus 0 day neonate lung cDNA, RIKEN full-length enriched  
library, clone:B030025C01 product:similar to CYTOCHROME P450 2S1  
[Homo sapiens], full insert sequence.









```

Db      1469 GGACTTTTCAACATCCCGCGGACTTCCAGTCCGGTCTGGGCCACT 1516
|||||
RESULT 3
BU552396
LOCUS   BU552396          900 bp      mRNA      linear      EST 16-SEP-2002
DEFINITION AGENCOURT_10333010 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6576174
5', mRNA sequence.
ACCESSION BU552396
VERSION   BU552396.1 GI:22902668
KEYWORDS EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 900)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTP
cDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM2774 row: e column: 06
High quality sequence stop: 800.
FEATURES             location/Qualifiers
source               1..900
                    /organism="Homo sapiens"
                    /mol_type="mRNA"
                    /db_xref="taxon:9606"
                    /clone="IMAGE:6576174"
                    /tissue_type="carcinoma, cell line"
                    /lab_host="DH10B (phage-resistant)"
                    /clone_lib="NIH_MGC 40"
                    /note="Organ: prostate; vector: pOTB7; site_1: XhoI;
                    site_2: EcoRI; cDNA made by oligo-dr priming.
                    Directionally cloned into EcoRI/XhoI sites using the
                    following 5' adaptor: GGCACGAG(G). Library constructed by
                    Ling Hong in the laboratory of Gerald M. Rubin (University
                    of California, Berkeley) using ZAP-cDNA synthesis kit
                    (Stratagene) and Superscript II RT (Life Technologies).
                    Note: this is a NIH_MGC Library."
ORIGIN
Alignment Scores:
Pred. No.:      1..99e-149      Length:      900
Score:          1529.00      Matches:      298
Percent Similarity: 99.67%      Conservative: 1
Best Local Similarity: 99.33%      Mismatches: 1
Query Match:     58.47%      Indels:      1
DB:              13      Gaps:          0

US-10-669-693-2 (1-504) x BU552396 (1-900)

QY      115 GlyValPhePheSerAsnGlyGluAArgGlnLeuAArgLysPheThrMetLeuAla 134
|||||
Db      2 GGGGTTTCTTCTCCAC-GGGGAGCGGTGAGGAGCTGAGGAAGTTTACCATGCTGCT 60
|||||
QY      135 LeuArgAspLeuGlyMetGlyLysArgGluGlyGluLeuLeuGlnAlaGluAlaArg 154
|||||
Db      61 CTGCGGAGCTGGGCACTGGGAGCCGAGAGGAGCGAGAGCTGATCCAGGCGGCGCCG 120
|||||
QY      155 CysLeuValGluThrPheGlnGlyThrGluGlyArgProPheAspProSerLeuLeu 174
|||||
Db      121 TGCTGCTGGAGACATTCACAGGGGACAGAGAGCGCCCATTCATCCCTCCCTGCTGCTG 180
|||||
QY      175 AlaGlnAlaThrSerAsnValValCysSerLeuLeuPheGlyLeuAArgPheSerTyrGlu 194
|||||

```

```

Db      181 GCCAGGCGACCTCCAACTAGTCTCTCTCCCTCTCTTTGGCTCGCTTCTCTATGAG 240
|||||
QY      195 AspLysGluPheGlnAlaValValArgAlaAlaGlyGlyThrLeuLeuGlyValSerSer 214
|||||
Db      241 GATAAGGAGTCTCCAGGCGCTGGTCCGGGAGCTGGTGGTACCTGCTGGGAGTACGATCC 300
|||||
QY      215 GlnGlyGlyGlnThrTyrGluMetPheSerTrpPheLeuAArgProLeuProGlyProHis 234
|||||
Db      301 CAGGGGGGTTCAGACCTACGAGATGTTCTCTGGTCTCTCGGCGCCCTGCCAGGCCCCAC 360
|||||
QY      235 LysGlnLeuLeuHisHisValSerThrLeuAlaAlaPheThrValArgGlnValGlnGln 254
|||||
Db      361 AAGCAGCTCTCTCCACAGCTCAGCCTTGGCTGCTTCCACATGCTCGGAGGTGCAGCAG 420
|||||
QY      255 HisGlnGlyAsnLeuAspAlaSerGlyProAlaAArgAspLeuValAspAlaPheLeuLeu 274
|||||
Db      421 CACCAGGGGAACCTGGATGCTTCGGGCGCCGACGTCGACCTTGTGATGCTTCTGCTG 480
|||||
QY      275 LysMetAlaGlnGluGlnAsnProGlyThrGluPheThrAsnLysAsnMetLeuMet 294
|||||
Db      481 AAGATGGCACAGGAGGAACAAACCCAGGCACAGAAATTCACCAACAGAACATGCTGATG 540
|||||
QY      295 ThrValIleTyrLeuLeuPheAlaGlyThrMetThrValSerThrThrValGlyTyrThr 314
|||||
Db      541 ACAGTCATTTATTTCTGTTTCTGGGACGATGACGGTCAGCACCGCTCGGCTATATACC 600
|||||
QY      315 LeuLeuLeuLeuMetLysTyrProHisValGlnLysTrpValAArgGluGluLeuAsnArg 334
|||||
Db      601 CTCCTGCTCTGATGAATACCTCATGTCCAAAGTGGTACGTGAGGAGCTGAATCGG 660
|||||
QY      335 GluLeuGlyAlaGlyGlnAlaProSerLeuGlyAspArgThrArgLeuProTyrThrAsp 354
|||||
Db      661 GAGCTGGGGCTGGCCAGGCACCAAGCCTAGGGGACGTACCCGCTCCCTTACACCGAC 720
|||||
QY      355 AlaValLeuHisGluAlaGlnArgLeuLeuAlaLeuValProMetGlyIleProArgThr 374
|||||
Db      721 GCGGTTCTGCATGAGGCGCGCGCTGCTGGCTGGTGGCCATGGGAATACCCGCGACC 780
|||||
QY      375 LeuMetArgThrThrArgPheArgGlyTyrThrLeuProGlnGlyThrGluValPhePro 394
|||||
Db      781 CTCATGCGGACACCCGCTCCGAGGGTACACCTGCCAGGACGAGGCTTCTCC 840
|||||
QY      395 LeuLeuGlySerIleLeuHisAspProAsnIlePheLysHisProGluGluPheAsnPro 414
|||||
Db      841 CTCCTGGCTCCATCTCTGATGATGACCCCAACATCTTTCAGCACCCAGAGAGTTCACCCC 900
|||||
RESULT 4
BU849902
LOCUS   BU849902          924 bp      mRNA      linear      EST 16-OCT-2002
DEFINITION AGENCOURT_10441031 NIH_MGC_109 Homo sapiens cDNA clone
IMAGE:6598606 5', mRNA sequence.
ACCESSION BU849902
VERSION   BU849902.1 GI:24034865
KEYWORDS EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 924)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM2831 row: k column: 22
High quality sequence stop: 701.

```



US-10-669-693-2 (1-504) x BQ900302 (1-903)

```

QY 115 GlyValPheSerAsnGlyGluArgTyrArgGlnLeuArgLysPheThrMetLeuAla 134
Db 2 GGGGTTTCTTCTCCAC -GGGAGCGGTGGAGCAGCTGAGAA-TTTACCATGCTGTCT 59
QY 135 LeuArgAspLeuGlyMetGlyLysArgGluGlyGluLeuLeuLeuAlaGluAlaArg 154
Db 60 CTGCGGACCTGGGATGGGAAGCAGAGGAGGAGCTGATCCAGCGGAGGCCCGG 119
QY 155 CysLeuValGlnThrPheGlnGlyThrGluGlyArgProPheAspProSerLeuLeu 174
Db 120 TGTCTGGTGGAGACATTCCAGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 179
QY 175 AlaGlnAlaThrSerAsnValValCysSerLeuLeuPheGlyLeuArgPheSerTyrGlu 194
Db 180 GCCCAGGACCTCCAAAGTAGTCTGCTCCCTCCCTCTTGGCTCCGCTTCTCTATGAG 239
QY 195 AspLysGluPheGlnAlaValValArgAlaAlaGlyGlyThrLeuLeuGlyValSer 214
Db 240 GATAAGGAGTTCAGGCGGTGTCGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 299
QY 215 GlnGlyGlyGlnThrTyrGlnMetPheSerTrpPheLeuArgProLeuProGlyProHis 234
Db 300 CAGGGGGTTCAGAGCTACGAGATGTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 359
QY 235 LysGlnLeuLeuHisHisValSerThrLeuAlaAlaPheThrValArgGlnValGln 254
Db 360 AAGCAGCTCTCCACAGCTCAGCACCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 419
QY 255 HisGlnGlyAsnLeuAspAlaSerGlyProAlaArgAspLeuValAlaPheLeuLeu 274
Db 420 CACGAGGGAACCTGGATGCTTTCGGGCGCCGACGTCGCTGCTGCTGCTGCTGCTGCTGCT 479
QY 275 LysMetAlaGlnGluGlnAsnProGlyThrGluPheThrAsnLysAsnMetLeuMet 294
Db 480 AAGATGCGCAGGAGGAGAACAAACCCAGGACAGAAATTCACCAACAGAACATGCTGATG 539
QY 295 ThrValIleTyrLeuLeuPheAlaGlyThrMetThrValSerThrValGlyTyrThr 314
Db 540 ACAGTCATTATTGCTGTTGCTGGAGAGATGACGGTCAGCAGCGGTGCGGTATACCC 599
QY 315 LeuLeuLeuMetLysTyrProHisValGlnLysTyrValArgGluGluLeuAsnArg 334
Db 600 CTCCTGCTCTGATGAATACCCCTCATGTCACAAAGTGGGTACGTGAGGAGCTGAATCGG 659
QY 335 GluLeuGlyAlaGlyGlnAlaProSerLeuLysAspArgThrArgLeuProTyrThrAsp 354
Db 660 GAGCTGGGGCTGGCCAGGACCAAGCCTAGGGGACCGTACCGGCTCCCTTACACCGAC 719
QY 355 AlaValLeuHisGluAlaGlnArgLeuLeuAlaLeuValProMetGlyIleProArgThr 374
Db 720 GCGGTTCTGATGANGCGCANGCGCTGCTGGCGCTGCTGGCGATGGGAATACCCCGCAC 779
QY 375 LeuMetArgThrThrArgPheArgGlyTyrThrLeuProGlnGlyThrGluValPhe-Pr 394
Db 780 CTATCGGACCAACCGCTTCCGAGGATACACCTGCCCCCANGGACCGGAGGCTTCCCC 839
QY 394 oLeuLeuGlySerIleLeuHisAspProAsnIlePheLysHisProGluGluPheAsnPr 414
Db 840 CCTCTTGGCTCCATCTGATGACCCCACTCTTCCAGCACCCAGAGAGTTCACACC 899
QY 414 oAsp 415
Db 900 AGAC 903

```

RESULT 6

BQ958112

LOCUS

DEFINITION AGENCOURT\_10034258 NIH\_MGC\_40 Homo sapiens cDNA clone IMAGE:6481143

5', mRNA sequence.

ACCESSION BQ958112

VERSION  
KEYWORDS

SOURCE  
ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BQ958112.1 GI:22373590

EST.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 939)

NIH-MGC http://mgi.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs@mail.nih.gov

Tissue Procurement: DCTD/DTF

cDNA Library Preparation: Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LNCM2662 row: a column: 16

High quality sequence stop: 636.

Location/Qualifiers

source

1..939

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:6481143"

/tissue\_type="carcinoma, cell line"

/lab\_host="DH10B (phage-resistant)"

/clone\_lib="NIH\_MGC\_40"

/note="Organ: prostate; Vector: pOTB7; Site 1: XhoI;

Site 2: EcoRI; cDNA made by oligo-dT priming.

Directionally cloned into EcoRI/XhoI sites using the

following 5' adaptor: GGCACGAG(G). Library constructed by

Ling Hong in the laboratory of Gerald M. Rubin (University

of California, Berkeley) using ZAP-cDNA synthesis kit

(Stratagene) and Superscript II RT (Life Technologies).

Note: this is a NIH\_MGC Library."

ORIGIN

Alignment Scores:

Pred. No.: 5,87e-137 Length: 939  
Score: 1410.50 Matches: 290  
Percent Similarity: 95.42% Conservative: 2  
Best Local Similarity: 94.77% Mismatches: 8  
Query Match: 53.94% Indels: 7  
DB: 13 Gaps: 1

US-10-669-693-2 (1-504) x BQ958112 (1-939)

```

QY 115 GlyValPheSerAsnGlyGluArgTyrArgGlnLeuArgLysPheThrMetLeuAla 134
Db 2 GGGGTTTCTTCTCCAC -GGGAGCGGTGGAGCAGCTGAGAA-TTTACCATGCTGTCT 60
QY 135 LeuArgAspLeuGlyMetGlyLysArgGluGlyGluLeuLeuGlnAlaGluAlaArg 154
Db 61 CTCGCGGACCTGGGCATGGGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
QY 155 CysLeuValGlnThrPheGlnGlyThrGluGlyArgProPheAspProSerLeuLeu 174
Db 121 TGTCTGGTGGAGACATTCCAGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
QY 175 AlaGlnAlaThrSerAsnValValCysSerLeuLeuPheGlyLeuArgPheSerTyrGlu 194
Db 181 GCCCAGGACCTCCAAAGTAGTCTGCTCCCTCCCTCTTTGGCTCCGCTTCTCTATGAG 240
QY 195 AspLysGluPheGlnAlaValValArgAlaAlaGlyGlyThrLeuLeuGlyValSer 214
Db 241 GATAAGGAGTTCAGGCGGTGCTCGGCGAGCTGGTGGTACCTGCTGGAGAGTCAGCTCC 300
QY 215 GlnGlyGlyGlnThrTyrGlnMetPheSerTrpPheLeuArgProLeuProGlyProHis 234
Db 301 CAGGGGGGTTCAGACCTACGAGATGTTCTCTGCTGCTTCTCGCGCCCTGCCAGGCCCCAC 360

```

```
Qy 235 LysGlnLeuLeuHisHisValSerThrLeuAlaLapheThrValArgGlnValGlnGln 254
Db 361 AAGCAGCTCTCCACACGTCAGCACCTTGGCTGCTTACAGTCCGCGAGTGCAGCAG 420
Qy 255 HisGlnGlyAsnLeuAspAlaSerGlyProAlaArgAspLeuValAspAlaPheLeuLeu 274
Db 421 CACCAGGGAACCTGGATGCTTCGGGGCCCGCAGCTGACCTTGTGTGATGCTTCTCTGTG 480
Qy 275 LysMetAlaGlnGlnGlnAsnProGlyThrGluPheThrAsnLysAsnMetLeuMet 294
Db 481 AAGATGGCAGAGGAGAAACCAACCCAGGCACAGATTCCACCAACAGACATGCTGATG 540
Qy 295 ThrValIleTyLeuLeuPheAlaGlyThrMetThrValSerThrThrValGlyTyThr 314
Db 541 ACAGTCAITTAITTTGCTGTTTCTGGGACGATGACGGTCAGCACCGTCCGGCTATACC 600
Qy 315 LeuLeuLeuLeuMetLysTyProHisValGlnLysTrpValArgGluLeuLeuAsnArg 334
Db 601 CTCCTGCTCTGATGAATACCTCATGTCCAAAAGTGGGTACGTGAGGAGCTGAATCGG 660
Qy 335 GluLeuGlyAlaGlyGlnAlaProSerLeuGlyAspArgThrArgLeuProTyThrAsp 354
Db 661 GAGCTGGGGGCTGGCCAGCACCAACAGCTAGGGGACCGTACCGGCTCTTACACCGAC 720
Qy 355 AlaValLeuHisGluAlaGlnArgLeuLeuAlaValProMetGlyIleProArgThr 374
Db 721 GCGGNTCTGCATGAGGCGCAGCCGCTGCTGGCGCTGGTCCCATGGGAATACCCCGACC 780
Qy 375 LeuMetArgThrThrArg-PheArgGlyTyThrLeuProGlnGly-ThrGlu---ValP 393
Db 781 CTCATCGGGACACCGCTTTCGAGGGTACACCTTGGCCCAAGNCACGGNAGGTCTTT 840
Qy 393 heProLeuLeuGlySer-IleLeuHis-AspProAsn-IlePheLysHisProGluGluP 412
Db 841 CCCCCNCTTGGCTCCACTCTGCATGGACCCCAACAATCTTCAAGCACCCCAAGAAGT 900
Qy 412 heAsnPro 414
Db 901 TTCAAGCC 908

RESULT 7
AK054324
LOCUS
DEFINITION
AK054324 2742 bp mRNA linear HTC 20-SEP-2003
Mus musculus 2 days pregnant adult female ovary cDNA, RIKEN
full-length enriched library, clone:E330015C15 product:similar to
CYTOCHROME P450 2S1 [Homo sapiens], full insert sequence.
ACCESSION AK054324
VERSION AK054324.1 GI:26096313
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
2
Genome Res. 10 (10), 1617-1630 (2000)
20499374
PUBMED 11042159
REFERENCE
3
Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
Konno,H., Akiyama,J., Nishi,K., Kitsuai,T., Tashiro,H., Itoh,M.,
Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
PUBMED 11076861
REFERENCE
4
The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 2742)
Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,M.,
Hayashida,K., Hayatsu,N., Hiranoto,K., Hiraoka,T., Hirozane,T.,
Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Tagami,M., Tégawa,K., Takahashi,F., Takaku-Akahira,S.,
Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
Muramatsu,M. and Hayashizaki,Y.
Direct Submission
Submitted (16-JUN-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL:http://genome.gsc.riken.go.jp/
URL:http://fantom.gsc.riken.go.jp/.
FEATURES
source
1 .2742
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:E330015C15"
/db_xref="MGI:2426416"
/db_xref="taxon:10090"
/clone="E330015C15"
/sex="female"
/tissue_type="ovary"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="2 days pregnant adult"
misc_feature
1 .2742
/note="similar to CYTOCHROME P450 2S1 [Homo sapiens]
(SPT9B266, evidence: FASTA, 78.4%ID, 99%length,
match=1500)"
ORIGIN
Alignment Scores:
Pred. No.: 1.24e-132 Length: 2742
Score: 1376.50 Matches: 268
Percent Similarity: 82.26% Conservative: 38
Best Local Similarity: 72.04% Mismatches: 43
Query Match: 52.64% Indels: 23
```



```
DB: 11 Gaps: 4
US-10-669-693-2 (1-504) x AK054324 (1-2742)

QY 60 LeuSerLysLysTyrGlyProValPheThrIleTyrLeuGlyProTyrArgProValVal 79
DB 33 CTAAGTAAGAAGATATGGGCTGTGTTCACGGTATACCTGGGCCCTTGGCGCGCGTGGTG 92
QY 80 ValLeuValGlyGlnGlnAlaValArgGluAlaLeuGlyGlyGlnAlaGluPheSer 99
DB 93 GTCCTGGTGGACATGATGCTGTAGAGAGAGCTTGGGAGGTGAGCTCAGAAATTCAGC 152
QY 100 GlyArgGlyThrValAlaMetLeuGluGlyThrPheAspGlyHisGlyValPhePheSer 119
DB 153 GGGCGTGAACATTCGCAACGCTGACAGAGCTTTGATGGTACCGAGTTTCTTGGC 212
QY 120 AsnGlyGluArgTyrArgGlnLeuArgLysPheThrMetLeuAlaLeuArgAspLeuGly 139
DB 213 AATGGGAGCGGTGAAACAGCTGAGGAAATTCACCTGCTCGCTACGGGACCTGGGC 272
QY 140 MetGlyLysArgGlyGlyGlnGlnLeuIleGlnAlaGluAlaArgCysLeuValGluThr 159
DB 273 ATGGCAGACGAGAGGAGGAGCTGATCCAGCGGAGGTGACAGAGTCTGGTGAGGCT 332
QY 160 PheGlnGlyThrGluGlyArgProPheAspProSerLeuLeuLeuAlaGlnAlaThrSer 179
DB 333 TTCAGAGAGACAGAGGAGCTCCATTCAACCTTCCATGCTGCTGCTGCGCCAGGCCACTCT 392
QY 180 AsnValValCysSerLeuLeuPheGlyLeuArgPheSerTyrGluAspLysGluPheGln 199
DB 393 AATGTGCTGCTGCTGCTGCTTGGCATCGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCT 452
QY 200 AlaValValArgAlaAlaGlyGlyThrLeuLeuGlyValSerSerGlnGlyGlnThr 219
DB 453 GCTGTGATCCAGGAGCAGCAAGTGTGCTTGTAGGATCATGCTCTCCATGGGCGCAGGCC 512
QY 220 TyrGluMetPheSerTyrPheLeuArgProLeuProGlyProHisLysGlnLeuHis 239
DB 513 TACAGATGTTCTCTGCTACTGCAGCCCTGCGAGGCCCCACACACAGCTCCAGCAC 572
QY 240 HisValSerThrLeuAlaAlaPheThrValArgGlnValGlnGlnHisGlnGlyAsnLeu 259
DB 573 CACTGGGACCCCTGGCTGCTTCCATCATCCAGCAGGTACAGAAACACCCAGGAGCGCTTC 632
QY 260 AspAlaSerGlyProAlaArgAspLeuValAspAlaPheLeuLeuLysMetAlaGlnGlu 279
DB 633 CAACCTCAGGTCTGACGCTGATGCTGTGACCCCTTCTGCTAAAGATGGCAGGAG 692
QY 280 GluGlnAsnProGlyThrGluPheThrAsnLysAsnMetLeuMetThrValIleTyrLeu 299
DB 693 AAACAGACCCAGGTACAGAAATTCACCGAAGAACTTGTGATGACGGTCAATACCTG 752
QY 300 LeuPheAlaGlyThrMetThrValSerThrThrValGlyTyrThrLeuLeuLeuMet 319
DB 753 CTGTTTGTGGACCATGACCATCGGTGCGCCATCGCTATGCTGCTGCTGCTGCTGCTG 812
QY 320 LysTyrProHisValGlnLysTyrValArgGluGluLeuAsnArgGluLeuGlyAlaGly 339
DB 813 AGATACCTCAATCCAGCAGCGCTGCGGAGAGAGCTCATACAGAGCTGGGTCTGGC 872
QY 340 GlnAlaProSerLeuGlyAspArgThrArgLeuProTyrThrAspAlaValLeuHisGlu 359
DB 873 AGGGCTCCAAAGTCTCAGCGATGAGTGGCTCCCTTACAGGATGCGCTTTTACAGGAG 932
QY 360 AlaGlnArgLeuAlaLeuValProMetGlyIleProArgThrLeuMetArgThrThr 379
DB 933 GCACAGCGCTCTGGCATGTGTAACCATGGGATGCGCCACACCATCATCAGAGGACCACT 992
QY 380 ArgPheArgGlyTyrThrLeuProGln----- 388
DB 993 TGCTTCGAGGGTACACTCTGCCCAAGGTAGGTGATGGCAGCCAGAGCAATATCTC 1052
QY 389 -----GlyThrGluValPheProLeuLeuGlySerIleLeuHisAspProAsnIlePhe 406

Db 1053 TGCCCTGGGACTTGGCTT-----TTGATATCCCTGTCTCTGGAGTACCT----- 1097
QY 407 LysHisProGlu-----GluPheAsnPro 414
Db 1098 ---CACCCCAACCACTCCACCTTCTTCTCTCC 1130

RESULT 8
BM047426 857 bp mRNA linear EST 07-NOV-2001
603628712F1 NTH_MGC_40 Homo sapiens cDNA clone IMAGE:5457075 5',
LOCUS mRNA sequence.
DEFINITION BM047426
ACCESSION BM047426
VERSION 1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 857)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/DTF
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LICM1956 row: h column: 04
High quality sequence stop: 817.
FEATURES
Location/Qualifiers
1..857
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5457075"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_40"
/notes="Organ: prostate; Vector: pOTB7; Site 1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCAGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

ORIGIN
Alignment Scores:
Pred. No.: 6,26e-132 Length: 857
Score: 1362.00 Matches: 277
Percent Similarity: 97.54% Conservative: 0
Best Local Similarity: 97.54% Mismatches: 7
Query Match: 52.08% Indels: 3
DB: 12 Gaps: 0
US-10-669-693-2 (1-504) x BM047426 (1-857)

QY 114 HisGlyValPhePheSerAsnGlyGluArgTyrArgGlnLeuArgLysPheThrMetLeu 133
Db 3 CATGGGTTTCTTCTCCAACGGGAGCGGTGGAGCAGCTGAGGAAGTTTACATGCTT 62
QY 134 AlaLeuArgAspLeuGlyMetGlyLysArgGluGlyGluGluLeuIleGlnAlaGluAla 153
Db 63 GCTCTCGGGACCTGGCATGGGAGCGAGAGCGAGAGCTGATCCAGCGGAGGCC 122
QY 154 ArgCysLeuValGluThrPheGlnGlyThrGluGlyArgProPheAspProSerLeuLeu 173
```

123 CGGTGTCGTGGAGACATTCAC- GGGACAGAGAGACGCCCATTCGATCCCTCCCTGCTG 181  
Db  
174 LeuAlaGlnAlaThrSerAsnValValCysSerLeuLeuPheGlyLeuArgPheSerTyr 193  
Qy  
182 CTGGCCAGGACACTCAAGTAGTGTCTCCCTCTCTTTGGCTCGCTTCTCTAT 241  
Db  
194 GluAspLysGluPheGlnAlaValValArgAlaAlaGlyGlyThrLeuLeuGlyValSer 213  
Qy  
242 GAGGATAAGGAGTTCAGGCGGTGTCGGGAGCTGGTGTGTACCTCTCGGAGTCAGC 301  
Db  
214 SerGlnGlyGlnThrTyrGluMetPheSerTrpPheLeuArgProLeuProGlyPro 233  
Qy  
302 TCCAGGGGGGFCAGACCTACAGAGTGTCTCTGTTCTGCTCCGCGCTGCCAGGCC 361  
Db  
234 HisLysGlnLeuLeuHisHisValSerThrLeuAlaAlaPheThrValArgGlnValGln 253  
Qy  
362 CACAGCAGCTCTCCACACAGTCAGCCTTGCTGCTTCTGCTCCGCGCTGCCAGGCC 421  
Db  
254 GlnHisGlnGlyAsnLeuAspAlaSerGlyProAlaArgAspLeuValAspAlaPheLeu 273  
Qy  
422 CAGCACAGGGGAACCTGGATGCTTTCGGGCCCCGACGTCGCTTGTGATGCCCTTCCTG 481  
Db  
274 LeuLysMetAlaGlnGluGlnAsnProGlyThrGluPheThrAsnLysAsnMetLeu 293  
Qy  
482 CTGAAGATGGCAGAGGAGCAAAACCCAGGCACAGATTTCACCAAGAACATGCTG 541  
Db  
294 MetThrValIleThrLeuLeuPheAlaGlyThrMetThrValSerThrThrValGlyTyr 313  
Qy  
542 ATGACAGTATTATTGCTGTTGCTGGAGCATGACGTCAGCACACGCTCGGCTAT 601  
Db  
314 ThrLeuLeuLeuMetLysTyrProHisValGlnLysTrpValArgGluGluLeuAsn 333  
Qy  
602 ACCCTCCTCTCTGATGAATACCCCTCATGTCCAAAGTGGGTACGTGAGGAGCTGAAT 561  
Db  
334 ArgGluLeuGlyAlaGlyGlnAlaProSerLeuGlyAspArgThrArgLeuProTyrThr 353  
Qy  
662 CGGAGCTGGGGCTGGCCAGGACCAAGCTAGGGAGCGTACCCGCTCCCTTACACC 721  
Db  
354 AspAlaValLeuHisGluAlaGlnArgLeuLeuAlaLeuValProMetGlyIleProArg 373  
Qy  
722 GAGCGGT-CTGCATGAGGCGCAGCGTCTGCTGGCGTGTGTGCCATGGGAAATACCCCG 780  
Db  
374 ThrLeuMetArgThrThrArgPheArgGlyTyrThrLeuProGlnGlyThrGluValPhe 393  
Qy  
781 AACCTCATGCGGACACCCGTTTCCGAGGGTACACCTG-CCCCAGGCCACGAGGCTTC 839  
Db  
394 ProLeuLeuGly 397  
Qy  
840 CCCCTCCTGGGT 851  
Db

RESULT 9  
BU543186  
LOCUS  
DEFINITION BU543186 877 bp mRNA linear EST 13-SEP-2002  
AGENCOURT\_10338701 NIH\_MGC\_40 Homo sapiens cDNA clone IMAGE:6575194  
S', mRNA sequence.  
ACCESSION BU543186  
VERSION BU543186.1 GI:22853669  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 877)  
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: DCTD/DFP  
cDNA Library Preparation: Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: L1CM2771 row: 1 column: 10  
High quality sequence stop: 703.

FEATURES

source

1. 877  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6575194"  
/tissue\_type="carcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_40"  
/note="Organ: prostate; Vector: pOTB7; Site 1: XhoI;  
Site 2: EcoRI; cDNA made by oligo-dT priming.  
Directionally cloned into EcoRI/XhoI sites using the  
following 5' adaptor: GGCACGAG(G). Library constructed by  
Ling Hong in the laboratory of Gerald M. Rubin (University  
of California, Berkeley) using ZAP-cDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies).  
Note: this is a NIH\_MGC Library."

ORIGIN

Alignment Scores:  
Pred. NO.: 1.31e-127 Length: 877  
Score: 1321.00 Matches: 271  
Percent Similarity: 94.46% Conservative: 2  
Best Local Similarity: 93.77% Mismatches: 10  
Query Match: 50.52% Indels: 7  
DB: 13 Gaps: 2

US-10-669-693-2 (1-504) x BU543186 (1-877)

Qy 115 GlyValPhePheSerAsnGlyGluArgTrpArgGlnLeuArgLysPheThrMetLeuAla 134  
Db 2 GGGGTTTCTTCTCCAC-GGGAGCGGTGGAGCAGCTGAGGAAGTTTACCATGCTGCT 60  
Qy 135 LeuArgAspLeuGlyMetGlyLysArgGluGlyGluLeuLeuGlnAlaGluAlaArg 154  
Db 61 CTCGGGACCTGGGCATGGGAAGCAGAGGAGGAGCTGATCCAGGCGGAGGCCCG 120  
Qy 155 CysLeuValGluThrPheGlnGlyThrGluGlyArgProPheAspProSerLeuLeu 174  
Db 121 TGTCTGTGTGAGACATTCAGGGGACAGAGGAGCCCATTCGATCCCTCCCTGCTGCTG 180  
Qy 175 AlaGlnAlaThrSerAsnValValCysSerLeuLeuPheGlyLeuArgPheSerTyrGlu 194  
Db 181 GCCCAGGCCACCTCCCAACGTAGTCTGCTCCTCTCTTTGGCCTCCGCTTCTCTATGAG 240  
Qy 195 AspLysGluPheGlnAlaValVala: gAlaAlaGlyGlyThrLeuLeuGlyValSerSer 214  
Db 241 GATAAGGAGTTCCAGGCGGTGGTCCGGGACGCTGGTGTACCTGCTGGGAGTCAGCTCC 300  
Qy 215 GlnGlyGlyGlnThrTyrGluMetPheSerTrpPheLeuArgProLeuProGlyProHis 234  
Db 301 CAGGGGGGTGAGACCTACGAGATGTTCTCTGTTCTGCGGCCCTTGCAGGCCCCAC 360  
Qy 235 LysGlnLeuLeuHisHisValSerThrLeuAlaAlaPheThrValArgGlnValGlnGln 254  
Db 361 AAGCAGCTCTCCACACGTCAGCACCTTGCTGCTTCCCTTCCAGTCCGCGAGTCAGCAG 420  
Qy 255 HisGlnGlyAsnLeuAspAlaSerGlyProAlaArgAspLeuValAspAlaPheLeuLeu 274  
Db 421 CACCAGGGGAACCTGGATGCTTCCGGGCCCGCAGCTGACCTTGTGAGAGCTTCTCTGCTG 480  
Qy 275 LysMetAlaGlnGluGlnAsnProGlyThrGluPheThrAsnLysAsnMetLeuMet 294  
Db 481 AAGATGGCAGAGGAGGAAACAAACCCAGGCACAGATTTCACCAACAGAACATGCTGATG 540  
Qy 295 ThrValIleTyrLeuLeuPheAlaGlyThrMetThrValSerThrThrValGlyTyrThr 314  
Db 541 ACAGTCATTATTGCTGTTTGTGCTGGGACGATGACGGTCCAGCACCCAGCGTATACCC 600



```

Qy 29 ArgGlyHisLeuProGlyProThrProLeuLeuGlyAsnLeuLeuGlnLeu 48
Db 106 CGTGCAACTTCCACACGAGCCCGTCCCTACCCCTCTTGGGGAACCTCTTGAGATG 165
Qy 49 ArgProGlyAlaLeuTyrSerGlyLeuMetArgLeuSerLysLeuTyrGlyProValPhe 68
Db 166 GACAGAGAGGCGCTCTCAAGTCTTTTATTTCAGCTTCGAGAAATATATGGAGATGTGTC 225
Qy 69 ThrIleTyrLeuGlyProTyrArgProValValLeuValGlyGlnGluAlaValArg 88
Db 226 ACAGTGCACCTGGGACCA---AGGCTGTGGTATGCTGTGGAAACAGACACCAATAGG 282
Qy 89 GluAlaLeuGlyGlyGlnAlaGluPheSerGlyArgGlyThrValAlaMetLeuGlu 108
Db 283 GAGGCTCTGTGGCCCAAGCCGAGGCTTCTCTGGCCGCGGACAGATTGCTGTCTGTGAG 342
Qy 109 GlyThrPheAspGlyHisGlyValPhePheSerAsnGlyGluArgTyrArgGlnLeuArg 128
Db 343 CCAACCTTCAAGGAATATGTTGATCTTTGCAATGGGAAACGTTGGAGACCTTCTGT 402
Qy 129 LysPheThrMetLeuAlaLeuArgAspLeuGlyMetGlyLysArgGluGlyGluLeu 148
Db 403 AGATTCTCTCGCCACCATGAGAGACTTGGGATGGGAAAGAGAGTGTGGAGGAGCGG 462
Qy 149 IleGlnAlaGluAlaArgCysLeuValGluThrPheGlnGlyThrGluGlyArgProPhe 168
Db 463 ATTGAGGAGAAAGCCCAATGTTTAGTGGAGAACTCGCGAAATCCGAGGAGCCCCCTG 522
Qy 169 AspProSerLeuLeuLeuAlaGlnAlaThrSerAsnValValCysSerLeuPheGly 188
Db 523 GACCCCAAGTCTCTTCAGTGCATCAGCCGCTTGGATGGGAAAGAGAGTGTGGAGGAGCGG 462
Qy 189 LeuArgPheSerTyrGluAspLysGluPheGlnAlaValAlaValArgAlaGlyThr 208
Db 583 GACCGTTTCAGTACACAGACCTGAGTCTTGGCCTGTGGAGTGTCTTATCAGACG 642
Qy 209 LeuLeuGlyValSerSerGlnGlyGlyGlnThrTyrGluMetPheSerTyrPheLeuArg 228
Db 643 TTTTCACTCATAGCTCATCTCCAGCCAGATGTTTGAGCTCTTCTGGCTCTCTCGAAG 702
Qy 229 ProLeuProGlyProHisLysGlnLeuLeuHisValSerThrLeuAlaAlaPheThr 248
Db 703 TACTTTCTGTGCCCCACAGCAAAATCTCAAAACCTGCAGAACTCTCGATACATT 762
Qy 249 ValArgGlnValGlnGlnHisGlnGlyAsnLeuAspAlaSerGlyProAlaArgAspLeu 268
Db 763 GGCCATAGTGTGGAGAGACAGGGCCACCTTGGACCCAGTGTCCA---CGAGACTTC 819
Qy 269 ValAspAlaPheLeuLeuLysMetAlaGlnGluGlnAsnProGlyThrGluPheThr 288
Db 820 ATTGATATTTACCTTCTCGCATGGAGAGGAAAGTCCAAACCCAGCACACGAGTTCAT 879
Qy 289 AsnLysAsnMetLeuMetThrValIleTyrLeuLeuPheAlaGlyThrMetThrValSer 308
Db 880 CACCAAGACCTTCATGATGTGTGCTCTCTCTCTTCTTGTGGCCACGAGACGAGCAGC 939
Qy 309 ThrThrValGlyTyrThrLeuLeuLeuMetLysTyrProHisValGlnLysTyrVal 328
Db 940 ACCAGCTCCGCTATGGCTCTCTCTCATGTCTAAGTACCCCAATGTTGCAGAGAAATC 999
Qy 329 ArgGluGlnLeuAsnArgGluLeuGlyValAlaGlyGlnAlaProSerLeuGlyAspArgThr 348
Db 1000 CAAAAGGAGATGTATGAGTATCGGTCTACACCGGTACCAACCCCTGATGACCGGACC 1059
Qy 349 ArgLeuProTyrThrAspAlaValLeuHisGlnAlaGlnArgLeuLeuAlaLeuValPro 368
Db 1060 AAAATGCCATACATGATGAGTATCATCCAGAGATTCAGAGATTTTCAGATCTTATACCT 1119
Qy 369 MetGlyIleProArgThrLeuMetArgThrThrArgPheArgGlyTyrThrLeuProGln 388
Db 1120 ATTGGAGTGCCACACAGAGTGTACCAAGATACCATGTTCCGAGGGTACCTGCTCCCAAG 1179
Qy 389 GlyThrGluValPheProLeuLeuGlySerIleLeuHisAspProAsnIlePheLysHis 408

```

```

Db 1180 AACACTGAGGTGTACCCCATCTCTGAGTTCAGCTCTTACATGATCCACAGTACTTTGAA 1239
Qy 409 ProGluGluPheAsnProAspArgPheLeuAspAlaAspGlyArgPheArgLysHisGlu 428
Db 1240 CCAGACAGTTTCAATCTCTGACCATCTCTGATGCAATGGGCGACTGAAGAAAGTGA 1299
Qy 429 AlaPheLeuProPheSerLeuGlyLysArgValCysLeuGlyGluGlyLeuAlaLysAla 448
Db 1300 GCTTTTCTGCGCTTCTCAACAGGAAGGCATTTGTCTTGTGTGAAGCATTTGCCCGCAAC 1359
Qy 449 GluLeuPheLeuPhePheThrThrIleLeuGlnAlaPheSerLeuGluSerProCysPro 468
Db 1360 GAGTGTGTTCTTCTTCTCAGTCCATCTCTCCAGAACTTCTCTGTGGCAGCCATGTGCT 1419
Qy 469 ProAspThrLeuSerLeuLysProThrValSerGlyLeuPheAsnIleProProAlaPhe 488
Db 1420 CTAAGGACATTTGACCTCTCTCCCAAGAGAGAGTGTGTGGAAATATCTCCACGTAC 1479
Qy 489 GlnLeu-----GlnValArgProThrAspLeu 497
Db 1480 CAGATCTGCTTCTTGGCCGCTGATTTGGCTGAGGACAGACAGTGGCCCGCAGTACTGTTG 1539
Qy 498 HisSerThr 500
Db 1540 AGAATGACT 1548

```

RESULT 11  
 BQ957100  
 LOCUS  
 DEFINITION 838 bp mRNA linear EST 21-AUG-2002  
 5', mRNA sequence.  
 AGENCOURT 8777144 NIH\_MGC\_40 Homo sapiens cDNA clone IMAGE:6384722  
 BQ957100  
 BQ957100.1 GI:22372578  
 EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 838)  
 NIH-MGC <http://mgc.nci.nih.gov/>  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)  
 Tissue Procurement: DCTD/DTP  
 CDNA Library Preparation: Rubin Laboratory  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: LHC2580 row: d column: 03  
 High quality sequence stop: 671.  
 Location/Qualifiers  
 1..838  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6384722"  
 /tissue\_type="carcinoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH\_MGC\_40"  
 /note="Organ: prostate; Vector: pOTB7; Site 1: XhoI;  
 Site 2: EcoRI; cDNA made by oligo-dT priming.  
 Directionally cloned into EcoRI/XhoI sites using the  
 following 5' adaptor: GGCACGAG(G). Library constructed by  
 Ling Hong in the laboratory of Gerald M. Rubin (University  
 of California, Berkeley) using ZAP-cDNA synthesis kit  
 (Stratagene) and Superscript II RT (Life Technologies).  
 Note: this is a NIH\_MGC Library."

FEATURES  
 source

ORIGIN

## Alignment Scores:

Pred. No.: 2,67e-117 Length: 838  
 Score: 1222.50 Matches: 248  
 Percent Similarity: 95.02% Conservative: 0  
 Best Local Similarity: 95.02% Mismatches: 10  
 Query Match: 46.75% Indels: 4  
 DB: 13 Gaps: 3

US-10-669-693-2 (1-504) x BQ957100 (1-838)

```

QY 115 GlyValPhePheSerAsnGlyGluArgTrrArgGlnLeuArgGlyPheThrMetLeuAla 134
DB 2 GGGGTTTTCTTCCAC-GGGGAGCGGTGGAGCGACCTGAGGAAGTTTACCATGCTTGCT 60

QY 135 LeuArgAspLeuGlyMetGlyLysArgGlyGluGluLeuLeuGlnAlaGluAlaArg 154
DB 61 CTGCGGGACCTGGGATGGGAAGCGAGGAGCGAGCTATCCAGCGGAGGCGCGG 120

QY 155 CysLeuValGluThrPheGlnGlyThrGluGlyArgPropheAspProSerLeuLeuLeu 174
DB 121 TGTCTGGTGGACATTCAGGGGACAGAGACGCCCATTCATCCCTCCCTGCTGCTG 180

QY 175 AlaGlnAlaThrSerAsnValValCysSerLeuLeuPheGlyLeuArgPheSerTyrGlu 194
DB 181 GCCCAGGCGACCTCCAGCTAGTCTGCTCCCTCCCTTTGGCTCCGCTTCCTATGAG 240

QY 195 AspLysGluPheGlnAlaValValArgAlaAlaGlyGlyThrLeuLeuGlyValSerSer 214
DB 241 GATAGGAGTTCAGGCGCGTGGTCCGGGCGAGTGGTGTACCTCTCGGAGTCAGCTCC 300

QY 215 GlnGlyGlyGlnThrTyrGluMetPheSerTrpPheLeuArgProLeuProGlyProHis 234
DB 301 CAGGGGGTTCAGACTAGAGATGTTCTCTCTGTTCTCGGGCCCTCCAGGCCCCCAGC 360

QY 235 LysGlnLeuLeuHisHisValSerThrLeuAlaAlaPheThrValArgGlnValGlnGln 254
DB 361 AAGCAGCTCTCTCCACCGTCAGCACCTTGGCTGCTTTCACAGTCCGCGAGGTGCGAG 420

QY 255 HisGlnGlyAsnLeuAspAlaSerGlyProAlaArgAspLeuValAlaPheLeuLeu 274
DB 421 CACAGGGGAACTGGATGCTTCGGGCGCCGCGACGTGACCTTGTGTGATCGCTTCCTG 480

QY 275 LysMetAlaGlnGluGlnAsnProGlyThrGluPheThrAsnLysAsnMetLeuMet 294
DB 481 AAGATGGCACAGGAGGAACAAACCCAGGCACAGAAATTCACCAACAGAAATGCTGATG 540

QY 295 ThrValIleThrLeuLeuPheAlaGlyThrMetThrValSerThrThrValGlyTyrThr 314
DB 541 ACATGCATTTATTTGCTGTTTGTGGGACGATGACGGTCAGCACCCAGCGTATATACC 600

QY 315 LeuLeuLeuMetLysTyrProHisValGlnLysTrpValArgGluGluLeuAsnArg 334
DB 601 CTCCTGCTCTGTAGTAATACCCCTCATGTCCAAAGTGGGTACGTGAGGAGCTGAATCGG 660

QY 335 GluLeuGlyAlaGlyGlnAlaProSerLeuGlyAspArgThrArgLeuProTyrThrAsp 354
DB 661 GAGCTGGGGGTGCGCCAGGCACCAAGCTAGGGACCGTACCCGCCCTCCCTTACCCGGA 720

QY 355 ---AlaValLeuHisGluAlaGlnArgLeu---LeuAlaLeuValProMetGly---Ile 371
DB 721 CGCGGGTCTGTCATGAGGCGGCANCGGCTGGCTGGCGGCTGGTGGCCCATGGGGAATA 780

QY 372 Pro 372
DB 781 CCC 783

```

## RESULT 12

BC022141  
 LOCUS BC022141  
 DEFINITION Mus musculus cDNA sequence BC034834, mRNA (cDNA clone IMAGE:5101144), containing frame-shift errors.  
 ACCESSION BC022141  
 VERSION BC022141.1 GI:18381021

1944 bp mRNA linear HTC 19-NOV-2003  
 Mus musculus cDNA sequence BC034834, mRNA (cDNA clone IMAGE:5101144), containing frame-shift errors.

## KEYWORDS

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 1944)  
 AUTHORS Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,

Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,  
 Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,  
 Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,  
 Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,  
 Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,F.L.,  
 Scheet,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,  
 Carninci,P., Brange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,  
 Abramson,R.D., Mullaly,S.J., Bosak,S.A., McEwan,P.J.,  
 McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,  
 Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,  
 Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,  
 Fahey,J., Helton,E., Kettman,M., Madan,A., Rodrigues,S.,  
 Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,  
 Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,  
 Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,  
 Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalley,D.E.,  
 Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.  
 human and mouse cDNA sequences  
 Generation and initial analysis of more than 15,000 full-length  
 human and mouse cDNA sequences  
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)  
 22388257  
 12477932  
 2 (bases 1 to 1944)  
 Strausberg,R.  
 Direct Submission  
 Submitted (24-JAN-2002) National Institutes of Health, Mammalian  
 Gene Collection (MGC), Cancer Genomics Office, National Cancer  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
 USA  
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
 Contact: MGC help desk  
 Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
 Tissue Procurement: Jeffrey E. Green, M.D.  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Genome Sequence Centre,  
 BC Cancer Agency, Vancouver, BC, Canada  
[info@bcgsc.bc.ca](mailto:info@bcgsc.bc.ca)  
 Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,  
 Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,  
 Letticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo  
 Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven  
 Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline  
 Schein, Duane Smalilus, Michael Smith, Lorraine Spence, Jeff Stott,  
 Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,  
 George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found  
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Series: IRAC Plate: 57 Row: i Column: 5  
 This clone was selected for full length sequencing because it  
 passed the following selection criteria: Hexamer frequency ORF  
 analysis  
 This clone has the following problem: frame shifted.

FEATURES  
source

Location/Qualifiers  
 1..1944  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="FVB/N"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:5101144"  
 /tissue\_type="Liver, normal. 5 month old male mouse."  
 /clone\_lib="NCI CGAP\_lig"  
 /lab\_host="DH10B"  
 /note="Vector: pCMV-SPORT6"

## ORIGIN

Alignment Scores:	1.13e-108	Length:	1944
Pred. No.:	1146.00	Matches:	230
Score:	67.28%	Conservative:	101
Percent Similarity:	46.75%	Mismatches:	155
Best Local Similarity:	43.82%	Indels:	8
Query Match:	11	Gaps:	3
DB:			
US-10-669-693-2 (1-504) x BC022141 (1-1944)			
Qy	7	TripAlaLeuLeuAlaLeuAlaLeuLeuLeuLeuLeuThrLeuAlaLeu-----	23
Db	21	TGGGTCTCCCTACTCTTCTCTCTC-CTGGTTTGGTGATGTCCTCAGCCTCTCTTCAGTGT	79
Qy	24	---SerGlyThrArgAlaAraGlyHisLeuProProGlyProThrProLeuProLeuLeu	42
Db	80	TGGCAAAAAATCGCACATGGGGGAGAGATGGCCCCCTGGCCCACTCCCACTTCCCATATT	139
Qy	43	GlyAsnLeuLeuGlnLeuArgProGlyAlaLeuTyrSerGlyLeuMetArgLeuSerLys	62
Db	140	GGGAATATACATGCAATTGGACTTCAAGACATCCCTGTCATGCTTTCAGCTGGCCAAA	199
Qy	63	LysTyrGlyProValPheThrIleTyrLeuGlyProTyrArgProValValValLeuVal	82
Db	200	GAGTATGGCCCTGTTACACTCTGTACTTTGGATCCTGG---CCTACCGTGGTCTCTACAT	256
Qy	83	GlyGlnGluAlaValArgGluAlaLeuGlyGlyGlnAlaGluGluPheSerGlyArgGly	102
Db	257	GGCTATGATGTGGTGAAGGAGCGCTACTCAATCAAGGTGATGAGTTCTCTTGGGAGAGA	316
Qy	103	ThrValAlaMetLeuGluGlyThrPheAspGlyHisGlyValPhePheSerAsnGlyGlu	122
Db	317	CCTCTGCCCATTTATCGAGATGACCGAAGGAGACATGGAAATGT--TTAGTCAAGGAGAG	375
Qy	123	ArgTyrArgGlnLeuAraGlySerPheThrMetLeuAlaLeuAraGlyAspLeuGlyMetGlyLys	142
Db	376	AGGTGGAAACTCCTCGGAGCGTCTCCCTCATGACGCTGAAGAACTTCGGAATGGGAAG	435
Qy	143	ArgGluGlyGluGluLeuIleGlnAlaGluAlaAraGlyCysLeuValGluThrPheGlnGly	162
Db	436	AGAAAGCTAGAGGAGAGGGGTGCAGGAGGAAGCCGGTGCCTCGTGGAGAGTTACATAAA	495
Qy	163	ThrGluGlyAraProPheAspProSerLeuLeuLeuAlaGlnAlaThrSerAsnValVal	182
Db	496	ACAGAGCTACGACCCCTTTGACCCCACTTCATCTTGGCTGTGCTCCCTGCAACGTGATC	555
Qy	183	CysSerLeuLeuPheGlyLeuAraGlyPheSerTyrGluAspLysGluPheGlnAlaValVal	202
Db	556	TGCTCCATCTTTTCAACGAGCGATCCCATACATGACACAGACATTCCTCAACCTCATG	615
Qy	203	ArgAlaAlaGlyGlyThrLeuLeuGlyValSerSerGlnGlyGlyGlnThrTyrGluMet	222
Db	616	GACCTTTTAAATAAAAAATTTTACCAACTGAACTCTATATGGATCCAGATGTACACCTTA	675
Qy	223	PheSerTyrPheLeuArgProLeuProGlyProHisLysGlnLeuLeuHisValSer	242
Db	676	TGGCCGACGATCATGAAGTACATACCTCTGGAAGCATAGAAATTCATAAAAGGCTTGCT	735
Qy	243	ThrLeuAlaAlaPheThrValArgGlnValGlnGlnHisGlnGlyAsnLeuAspAlaSer	262
Db	736	GGTGTTAAAAAATTTCATCTCTAGAAAAAGTGAAGGAGCACACAGAGTCCCTGGACCCGCG	795
Qy	263	GlyProAlaAraGlyLeuValAspAlaPheLeuLeuLysMetAlaGlnGluGlnAsn	282
Db	796	AAACCTT---CGGGATTACATCGACTGTTTTCCTCAGCAAGATCGAGGAGGAAACAAAC	852
Qy	283	ProGlyThrGluPheThrAsnLysAsnMetLeuMetThrValIleTyrLeuLeuPheAla	302
Db	853	CTGAATCTGATTTTAACTGGAGAAATTTGGCCCATCTGTGGGTCTAACTTGTTTACGGCA	912
Qy	303	GlyThrMetThrValSerThrThrValGlyTyrThrThrLeuLeuLeuMetLysTyrPro	322







```

Db 1292 TACTTCATGACTTTCTCAGCAGGAAACGGGTTTGTGCTGAGAGGGCGCTGGCCGCATG 1351
Qy 449 GluLeuPheLeuPheThrThrIleLeuGlnIaPheSerLeuGluSerProCysPro 468
|||||
Db 1352 GAGCTGTTTCTAATCCTGACCACCACTTTTACAGAACCTTTAAACTGAAATCTCTGTTCCAC 1411
Qy 469 ProAspThrLeuSerLeuLeuProThrValSerGlyLeuPheAsnIleProProAlaPhe 488
|||||
Db 1412 CCAAAAGACATTGATATGATCCCATTTGTGAATGATTAATCACTCTCCACCCCAATAC 1471
Qy 489 GlnLeu 490
|||||
Db 1472 CAGCTC 1477

RESULT 15
AK008688
LOCUS
DEFINITION
Mus musculus adult male stomach cDNA, RIKEN full-length enriched
library, clone:2210009K14 product:similar to CYTOCHROME P450 2C18
(EC 1.14.14.1) (CYP1C18) (P450-6B/29C) [Homo sapiens], full insert
sequence.
ACCESSION
AK008688
VERSION
AK008688.1 GI:12843035
KEYWORDS
HTC; CAP trapper.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
1
AUTHORS
Carninci,P. and Hayashizaki,Y.
TITLE
High-efficiency full-length cDNA cloning
JOURNAL
Meth. Enzymol. 303, 19-44 (1999)
MEDLINE
99279253
PUBMED
10349636
REFERENCE
2
AUTHORS
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL
Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE
20499374
PUBMED
11042159
REFERENCE
3
AUTHORS
Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
Konno,H., Akiyama,J., Nishi,K., Kitsuai,T., Tashiro,H., Itoh,M.,
Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
TITLE
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
JOURNAL
Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE
20530913
PUBMED
11076861
REFERENCE
4
AUTHORS
The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
TITLE
Functional annotation of a full-length mouse cDNA collection
JOURNAL
Nature 409, 685-690 (2001)
REFERENCE
5
AUTHORS
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
TITLE
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
JOURNAL
Nature 420, 563-573 (2002)
REFERENCE
6
AUTHORS
Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Arai,A., Aono,H.,
Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Fukunishi,Y.,
Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hiramoto,K.,
Hiraoka,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Izawa,M.,
Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M.,
Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K.,

```

```

Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C.,
Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D.,
Shibata,K., Shibata,Y., Shingawa,A., Shiraki,T., Sogabe,Y.,
Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T.,
Tejima,Y., Toya,Y., Yamamura,T., Yasunishi,A., Yoshida,K.,
Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
Direct Submission
Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)
Please visit our web site (http://genome.gsc.riken.go.jp/) for
further details.

COMMENT
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues. First strand cDNA was primed with a primer
[5' GAGAGAGAGAGGATCCAAAGACTCTTTTITTTTNN 3'], cDNA was
prepared by using rehealose thermo-activated reverse transcriptase
and subsequently enriched for full-length by cap-trapper. Second
strand cDNA was prepared with the primer adapter of sequence [5'
GAGAGAGAGATTCTCGAGTTAATTAAATCCCCCCCCC 3']. cDNA was cleaved
with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI.
Host: SOUR.

FEATURES
Location/Qualifiers
1..1870
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:2210009K14"
/db_xref="MGI:1905087"
/db_xref="taxon:10090"
/clone="2210009K14"
/sex="male"
/tissue_type="stomach"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
<1..1427
/note="unnamed protein product; putative
similar to CYTOCHROME P450 2C18 [EC 1.14.14.1] (CYP1C18)
(P450-6B/29C) [Homo sapiens] (SWISSPROT|P33260, evidence:
PASTY, 75%ID, 96.3%length, match=1416)"
/codon_start=3
/protein_id="BAB25834.1"
/db_xref="GI:12843036"
/translation="PKLWRONSQRNLPFGPTPLPIIGNILOLDKDISILNFSKV
YGFVFTLYLGRNPAVLHGYSYAEKAEFTDHEEAGRGVFFVDFKFKNCGVFSSGR
TWKERRPFLMTLRNFGMRRSIEDRIQOEARCLVDEURKTGKPCDPTFLUGASCN
VICSIVQNFQDYKQDFELFDILNNEVLISFQICNNFNPAVIDYLPGRHKLH
KNFAPBYFLSKVQHESLDINNPRDFIDCFILKMEQKHNPTEFTCNLVFTAS
DLFAAGTBTSTTLRYSLLLKYPVETAKVQEEIDCVIGHRSPCMDRHSMPYTA
VLHRIQYLDLTLPLSHAVTRDYKREYLLPKGTITVIASTLVLDQDEPNPKFD
PSHFDREGKFKSDYFPFSTGKRICVGEGLARAELEFLTLTLQNFNLKSPVDLKD
LDITPVGANGFASVPPKQICFIPI"
1851..1856
/note="putative"
1870
/note="putative"

polyA_signal
polyA_site

ORIGIN
Alignment Scores:
Pred. No.: 1,28e+100 Length: 1870
Score: 1069.00 Matches: 213
Percent Similarity: 64.94% Conservatives: 87
Best Local Similarity: 46.10% Mismatches: 160
Query Match: 40.88% Indels: 2
DB: 11 Gaps: 2
US-10-669-693-2 (1-504) x AK008688 (1-1870)

```

```
QY 29 ArgGlyHisLeuProProGlyProThrProLeuProLeuLeuGlyAsnLeuLeuGlnLeu 48
Db 30 AGAAGAAACCTCTCTCTGCCCCACACCTCTCTCCCATCATTTGAAATATCTTTCAGCTA 89
QY 49 ArgProGlyAlaLeuTyrSerGlyLeuMetArgLeuSerLysLysTyrGlyProValPhe 68
Db 90 GATCTTAAGACATCAGCAATCTCTGAGGAATTTTCAAAAGTCTATNGCCCTGTGTTTC 149
QY 69 ThrIleTyrLeuGlyProTyrArgProValValLeuValGlnGlnAlaValArg 88
Db 150 ACTCTGTACTTGGGCAGG---AATCCTGCTGTGTGTATCATGATGAAGCTGTGAAA 206
QY 89 GluAlaLeuGlyGlyGlnAlaGluPheSerGlyArgGlyThrValAlaMetLeuGlu 108
Db 207 GAAGCCTTTACTGATCATGCGGAGGAGTTTGTGGAAGAGAGTTTTCAGTGTGTTGAT 266
QY 109 GlyThrPheAspGlyHisGlyValPhePheSerAsnGlyGluArgTyrArgGlnLeuArg 128
Db 267 AGCTTTAAGAAACATCGCGGGTGTGTTTCAGCAGTGGGAGGACCTGGAAAGAGATGAGA 326
QY 129 LysPheThrMetLeuAlaLeuArgAspLeuGlyMetGlyLysArgGluGlyGlnLeu 148
Db 327 CGCTTCTCACTCATCTCTCGGAAACTTTGGGATGGGTAGGAGGACATTGAGGACCGA 386
QY 149 IleGlnAlaGluAlaArgCysLeuValGluThrPheGlnGlyThrGluGlyArgProPhe 168
Db 387 ATTCAGAGAGGCTCGCTGCTGTGTGGATGAATTGAGGAAACCAAGGTTGAACCTGT 446
QY 169 AspProSerLeuLeuAlaGlnAlaThrSerAsnValValCysSerLeuLeuPheGly 188
Db 447 GATCCACCTTCATCTCGGCTGTGCTTCTTGCAATGTCACTGTCTCCATGTTTCCAG 506
QY 189 LeuArgPheSerTyrGluAspLysGluPheGlnAlaValValArgAlaAlaGlyThr 208
Db 507 AATCGTGTGATTATAAAGATCAGGAATCTTACCTTCTTAGACATCTTTAAATGAAAT 566
QY 209 LeuLeuGlyValSerSerGlnGlyGlnThrTyrGluMetPheSerTrpPheLeuArg 228
Db 567 GTGAGATTCTGAGTTCCCTTGGATTTCAGATTGGCAATTAATTTTCTGCTGTCTGAT 626
QY 229 ProLeuProGlyProHisLysGlnLeuLeuHisValSerThrLeuAlaAlaPheThr 248
Db 627 TATCTACCGAAGACACAGAAAATACATAAAATTTTGCCTTTCAGAACATTACTTT 686
QY 249 ValArgGlnValGlnGlnHisGlnGlyAsnLeuAspAlaSerGlyProAlaArgAspLeu 268
Db 687 TTGTCAAAAGTAAACAGCATCAAGAGTCATCGGATATTAAACAATCCT---CGGACTTT 743
QY 269 ValAspAlaPheLeuLeuLysMetAlaGlnGluGlnAsnProGlyThrGluPheThr 288
Db 744 ATTGATTGTTTCTGATCAAAATGGACAGGAAAGACACACCAAGACGGAGTTTACT 803
QY 289 AsnLysAsnMetLeuMetThrValIleTyrLeuLeuPheAlaGlyThrMetThrValSer 308
Db 804 TGTGAGAACTTGGTCTTCACAGCATCTGATCTGTTGCTGCTGGACAGAGACAAAGC 863
QY 309 ThrThrValGlyTyrThrIleuLeuLeuMetLysTyrProHisValGlnLysTrpVal 328
Db 864 ACAACCCCTGAGGACTCTCTGCTGCTCCTGCTCAAGTATCCAGAAGTCCAGCCAAAGTC 923
QY 329 ArgGluGluLeuAsnArgGluLeuValAlaGlyGlnAlaProSerLeuGlyAspArgThr 348
Db 924 CAGGAGAGATCGACTGTGTGATTGGCAGACACAGGAGCCCTGCGATGCGAGGACGCAT 983
QY 349 ArgLeuProTyrThrAspAlaValLeuHisGlnAlaGlnArgLeuLeuAlaValPro 368
Db 984 AGCATGCCCTACACTGATGCTGTGTGATGATCCAGATACATGATGATCTCTCTCC 1043
QY 369 MetGlyIleProArgThrLeuMetArgThrArgPheArgGlyTyrThrLeuProGln 388
Db 1044 ACCAGCCTTCCCATGCCGTTGACCCGTGACGTGAAATTCAGAGAGTATCTCATTTCCCAAG 1103
```

```
QY 389 GlyThrCluValPheProLeuLeuGlySerIleLeuHisaspProAsnIlePheLysHis 408
Db 1104 GGCACCACTGTCTATAGCATCACTACTTCTGTCTATATGATAAAGAAATTTCCCAAC 1163
QY 409 ProGluGluPheAsnProAspArgPheLeuAspAlaAspGlyArgPheArgLysHisGlu 428
Db 1164 CCAGAGAAGTTTGACCCCGACCTTCTGATGAGCGGTGGCAAGTTTAAGAAAAGTGAC 1223
QY 429 AlaPheLeuProPheSerLeuGlyLysArgValCysLeuGlyGluGlyLeuAlaLysAla 448
Db 1224 TACTTTTTCCTTCTCAACAGGAAAACGATTTGTGGGAGAGGCTCTGGCTCGAGCA 1283
QY 449 GluLeuPheLeuPheThrThrThrLeuGlnAlaPheSerLeuGluSerProCysPro 468
Db 1284 GAATTATTTTATTTCTGACCACCATTTTACAGAACTTTTAACTTGAAATCTCCAGTTGAT 1343
QY 469 ProAspThrLeuSerLeuLysProThrValSerGlyLeuPheAsnIleProProAlaPhe 488
Db 1344 CTAAAGACCTTGATACTACTCCAGTTGCCAATGGATTTCTTCTGTGCCCCCAAGTTC 1403
QY 489 GlnLeu 490
Db 1404 CAGATA 1409
```

Search completed: September 18, 2004, 04:33:32  
Job time : 3799 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: September 17, 2004, 06:17:00 ; Search time 11362.9 Seconds  
(without alignments)  
11447.105 Million cell updates/sec

Title: US-10-669-693-3\_COPY\_1980\_4980  
Perfect score: 3001  
Sequence: 1 cttctctctcctaccctcc.....atcctggtaacacagttaa 3001

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0  
Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*

- 1: gb.ba.\*
- 2: gb.htg.\*
- 3: gb.in.\*
- 4: gb.om.\*
- 5: gb.ov.\*
- 6: gb.pat.\*
- 7: gb.ph.\*
- 8: gb.pl.\*
- 9: gb.pr.\*
- 10: gb.ro.\*
- 11: gb.sts.\*
- 12: gb.sy.\*
- 13: gb.un.\*
- 14: gb.vi.\*
- 15: em.ba.\*
- 16: em.fun.\*
- 17: em.hum.\*
- 18: em.in.\*
- 19: em.mu.\*
- 20: em.om.\*
- 21: em.or.\*
- 22: em.ov.\*
- 23: em.pat.\*
- 24: em.ph.\*
- 25: em.pl.\*
- 26: em.ro.\*
- 27: em.sts.\*
- 28: em.un.\*
- 29: em.vi.\*
- 30: em.htg\_hum.\*
- 31: em.htg\_inv.\*
- 32: em.htg\_other.\*
- 33: em.htg\_mus.\*
- 34: em.htg\_pln.\*
- 35: em.htg\_rod.\*
- 36: em.htg\_man.\*
- 37: em.htg\_vrt.\*
- 38: em\_sy.\*
- 39: em.htgo\_hum.\*
- 40: em.htgo\_mus.\*
- 41: em.htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	2997.8	99.9	129402	9	AC011510	AC011510 Homo sapi
c	2997.8	99.9	173251	2	AC058798	AC058798 Homo sapi
c	479.8	16.0	321519	2	AL714004	AL714004 Homo sapi
4	276.4	9.2	9532	9	HS999D10	Z94802 Human DNA s
5	272.2	9.1	126053	9	AL359916	AL359916 Human DNA
6	263.8	8.8	179607	2	AC027272	AC027272 Homo sapi
7	259.2	8.6	188833	9	HS268H5	AL008718 Human DNA
8	258	8.6	39634	9	AC139315	AC139315 Homo sapi
c	258	8.6	40948	9	AC140106	AC140106 Homo sapi
c	258	8.6	41867	2	AC144507	AC144507 Homo sapi
c	258	8.6	153609	9	HS479617	AJ009617 Homo sapi
c	258	8.6	158103	9	AC007952	AC007952 Homo sapi
c	258	8.6	158103	9	AC007952	AC007952 Homo sapi
c	258	8.6	176584	2	AC062010	AC062010 Homo sapi
14	257	8.6	79528	6	BD176843	BD176843 A method
16	257	8.6	79528	9	HS466N1	Z97630 Human DNA s
17	256.8	8.6	176221	2	AC011111	AC011111 Homo sapi
18	256.2	8.5	128611	9	AC083849	AC083849 Homo sapi
19	256	8.5	128624	9	AC008896	AC008896 Homo sapi
20	254.8	8.5	165735	9	AC135893	AC135893 Homo sapi
c	254.8	8.5	167110	2	AC022621	AC022621 Homo sapi
c	254.8	8.5	170269	2	AC144530	AC144530 Homo sapi
c	254.6	8.5	129036	9	AC005378	AC005378 Homo sapi
c	254.2	8.5	205642	9	AC016586	AC016586 Homo sapi
c	254.2	8.5	154179	9	AC087533	AC087533 Homo sapi
c	254.2	8.5	184511	2	AC011721	AC011721 Homo sapi
c	254.2	8.5	205638	9	AC009634	AC009634 Homo sapi
c	253	8.4	40351	9	AC005620	AC005620 Homo sapi
c	253	8.4	152945	9	AC098475	AC098475 Homo sapi
c	253	8.4	174539	9	AC103558	AC103558 Homo sapi
c	252.8	8.4	151999	2	AC005973	AC005973 Homo sapi
c	252.6	8.4	74155	9	AC004656	AC004656 Homo sapi
c	252	8.4	225370	9	AL445222	AL445222 Human DNA
c	251.8	8.4	63605	9	AL662787	AL662787 Human DNA
c	251.8	8.4	142916	9	AL928921	AL928921 Human DNA
c	251	8.4	171395	9	AC136605	AC136605 Homo sapi
c	251	8.4	179443	9	AC093290	AC093290 Homo sapi
c	251	8.4	182048	2	AC093293	AC093293 Homo sapi
c	251	8.4	186323	9	AC136601	AC136601 Homo sapi
c	251	8.4	262628	2	AC136609	AC136609 Homo sapi
c	250.8	8.4	2260	9	AY342392	AY342392 Homo sapi
c	250.8	8.4	162495	9	AC009996	AC009996 Homo sapi
c	250	8.3	69705	2	AC142832_3	Continuation (4 of
c	249.6	8.3	119043	9	AC139451	AC139451 Homo sapi
c	249	8.3	72539	9	AC119396	AC119396 Homo sapi

ALIGNMENTS

RESULT 1  
AC011510  
LOCUS AC011510 129402 bp DNA linear PRI 23-OCT-2000  
DEFINITION Homo sapiens chromosome 19 clone CTD-2195B23, complete sequence.  
AC011510  
ACCESSION AC011510.7 GI:10947023  
VERSION  
KEYWORDS HTG.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 129402)  
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.  
TITLE Direct Submission  
JOURNAL Unpublished

REFERENCE  
AUTHORS  
TITLE  
JOURNAL

2 (bases 1 to 129402)  
DOE Joint Genome Institute.  
Direct Submission  
Submitted (07-OCT-1999) Production Sequencing Facility, DOE Joint  
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
3 (bases 1 to 129402)  
DOE Joint Genome Institute and Stanford Human Genome Center.  
Direct Submission  
Submitted (23-OCT-2000) DOE Joint Genome Institute, 2800 Mitchell  
Drive, Walnut Creek, CA 94598, USA  
On Oct 23, 2000 this sequence version replaced gi:9211205.  
Draft Sequence Produced by DOE Joint Genome Institute  
www.jgi.doe.gov  
Finishing Completed at Stanford Human Genome Center  
www.shgc.stanford.edu  
Quality: Phrap Quality >=40 99.8% of Sequence;  
Estimated Total Number of Errors is 0.7.

FEATURES  
source  
1. .129402  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="19"  
/clone="CTD-2195B23"

ORIGIN  
Query Match : 99.9%; Score 2997.8; DB 9; Length 129402;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2999; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 CTTTCTCTCCTACCCAGCTAGTAAGATACGACCGGTGTTCCACCATCTACCT 60  
DB 23073 CTTTCTCTCCTACCCAGCTAGTAAGATACGACCGGTGTTCCACCATCTACCT 23132  
QY 61 GGGACCCCTGGCGCCTGTGTGCTCTGTTGGCAGGAGGTGTGCGGAGGCCCTGGG 120  
DB 23133 GGGACCCCTGGCGCCTGTGTGCTCTGTTGGCAGGAGGTGTGCGGAGGCCCTGGG 23192  
QY 121 AGGTCAAGCTGAGGAGTTAGCGGCCGGGAAACCGTAGGATGCTGGAAGGAGCTTTGA 180  
DB 23193 AGGTCAAGCTGAGGAGTTAGCGGCCGGGAAACCGTAGGATGCTGGAAGGAGCTTTGA 23252  
QY 181 TGGCCATGTAGTCAAGGCTGTAGGCTCTAGGCTCACAGCTGCCACCACTTACTCG 240  
DB 23253 TGGCCATGTAGTCAAGGCTGTAGGCTCTAGGCTCACAGCTGCCACCACTTACTCG 23312  
QY 241 TGTGTACCTTTGCACATGGCTTAGTCCCTCTGTGCTCTCATCTGTCAAATGGAGTGATA 300  
DB 23313 TGTGTACCTTTGCACATGGCTTAGTCCCTCTGTGCTCTCATCTGTCAAATGGAGTGATA 23372  
QY 301 ACAGTCCCATCAGCCGGGTGCGTGGCTAGTGGCTGAAATCCAAACACTTTGGGAGGCG 360  
DB 23373 ACAGTCCCATCAGCCGGGTGCGTGGCTAGTGGCTGAAATCCAAACACTTTGGGAGGCG 23432  
QY 361 GAGGTGGGTGGATCACTTGAGGTGAGGAGTTGAGACCAAGCTGGCCAAACATGGTGAAC 420  
DB 23433 GAGGTGGGTGGATCACTTGAGGTGAGGAGTTGAGACCAAGCTGGCCAAACATGGTGAAC 23492  
QY 421 CTTGTCTCTACTAAATAATAAAATAGCTGGCATGGTGGTACCTGTAAATCCCA 480  
DB 23493 CTTGTCTCTACTAAATAATAAAATAGCTGGCATGGTGGTACCTGTAAATCCCA 23552  
QY 481 GATACCTGGGAGTTGAGCGGAGGATCGTTGACCCGGGAGGAGATGTTGCGATGA 540  
DB 23553 GATACCTGGGAGTTGAGCGGAGGATCGTTGACCCGGGAGGAGATGTTGCGATGA 23612  
QY 541 ACCAAGACTGTGCCACTCGACTCTGGGCAACAGAGTGAGCCCTCCATCTCAAACAA 600  
DB 23613 ACCAAGACTGTGCCACTCGACTCTGGGCAACAGAGTGAGCCCTCCATCTCAAACAA 23672  
QY 601 ACAACAAAGAGCAGTGCCCATCATGTAGATTTAGTGAAGGAGTGAAGCTTTG 660  
DB 23673 ACAACAAAGAGCAGTGCCCATCATGTAGATTTAGTGAAGGAGTGAAGCTTTG 23732

QY 661 TCGAAGTGAGCACTCAGTAAATCAACAGGTTGTAGTATCAGTGATACCATCAATGATCC 720  
DB 23733 TCGAAGTGAGCACTCAGTAAATCAACAGGTTGTAGTATCAGTGATACCATCAATGATCC 23792  
QY 721 AGCTAAAGCCCTGAGGTTTCAAGAGATCCGAGCGCTTTCAAGTGTCTGGGATTTGGT 780  
DB 23793 AGCTAAAGCCCTGAGGTTTCAAGAGATCCGAGCGCTTTCAAGTGTCTGGGATTTGGT 23852  
QY 781 GGGCAAGCCCTCGAATAATAGAAAACAGTTCTCTGTATTACAAACAGAAAACAGGAGGCCCA 840  
DB 23853 GGGCAAGCCCTCGAATAATAGAAAACAGTTCTCTGTATTACAAACAGAAAACAGGAGGCCCA 23912  
QY 841 TCGTGGTGTCTGCGAGGAACCTCAGTAGTAACCTAAGACACACCGGTGCTGTCTCCCGAGC 900  
DB 23913 TCGTGGTGTCTGCGAGGAACCTCAGTAGTAACCTAAGACACACCGGTGCTGTCTCCCGAGC 23972  
QY 901 GCACCTAGGCCAGTGGGAAACAGACTCACACACAGTCCAGCCAGAGTGTGTCAAGGC 960  
DB 23973 GCACCTAGGCCAGTGGGAAACAGACTCACACACAGTCCAGCCAGAGTGTGTCAAGGC 24032  
QY 961 CAAGATGGGAAACAGCGGAGAAAGGTCAAGGTGGATGGGAGGGGTCAAGGCAAGAG 1020  
DB 24033 CAAGATGGGAAACAGCGGAGAAAGGTCAAGGTGGATGGGAGGGGTCAAGGCAAGAG 24092  
QY 1021 GGGTCAGGCCAGGCTGAGGGAAGCCCTGGGACTGTAGGAATTTAGAGGAGTACCTGAC 1080  
DB 24093 GGGTCAGGCCAGGCTGAGGGAAGCCCTGGGACTGTAGGAATTTAGAGGAGTACCTGAC 24152  
QY 1081 CCGGCATGTTTGGTGGGAGATTCAGGAAGTCTTCTCGAAGAGAGGTGTCTGGAGCTG 1140  
DB 24153 CCGGCATGTTTGGTGGGAGATTCAGGAAGTCTTCTCGAAGAGAGGTGTCTGGAGCTG 24212  
QY 1141 AGACTCATAAGATGAGTGGGAGGTGTTCCAGGCAGAAAGACCAAGCACTCAAAAGCA 1200  
DB 24213 AGACTCATAAGATGAGTGGGAGGTGTTCCAGGCAGAAAGACCAAGCACTCAAAAGCA 24272  
QY 1201 TGACTTTGAGAGAGCATTCATCCATTCAGTGAATTTTCAGACTGGGCAAGCTGGC 1260  
DB 24273 TGACTTTGAGAGAGCATTCATCCATTCAGTGAATTTTCAGACTGGGCAAGCTGGC 24332  
QY 1261 TCATGCTCTGTAATCCAGCACTTTGGAAGGCTGAATGGGAGGATGACTTTGAGCCCTAGGC 1320  
DB 24333 TCATGCTCTGTAATCCAGCACTTTGGAAGGCTGAATGGGAGGATGACTTTGAGCCCTAGGC 24392  
QY 1321 ATTGTGACAGCTGGGCAACATGGTGAAGACCTGCTCCACAAACAAACAAACAAAC 1380  
DB 24393 ATTGTGACAGCTGGGCAACATGGTGAAGACCTGCTCCACAAACAAACAAACAAAC 24452  
QY 1381 AAAAAATCAATTATACCTGTTACCTGAGGAGGCGGCTTATCTGCACTAAGACAATAGAGGAG 1440  
DB 24453 AAAAAATCAATTATACCTGTTACCTGAGGAGGCGGCTTATCTGCACTAAGACAATAGAGGAG 1500  
QY 1441 TGGTGTCTCTCTCTCTGTTGGGAGGCGGCTTATCTGCACTAAGACAATAGAGGAG 1560  
DB 24513 TGGTGTCTCTCTCTCTGTTGGGAGGCGGCTTATCTGCACTAAGACAATAGAGGAG 24572  
QY 1501 GGAATATAATCTCTCTCTGTTGAGGTTAGAGCAACACAGGCGCAAGGCATAT 1560  
DB 24573 GGAATATAATCTCTCTCTGTTGAGGTTAGAGCAACACAGGCGCAAGGCATAT 24632  
QY 1561 GTACGAGGTTAAGAGGGAATCAGGCAAGGCTTCTCAGAGAGGTGACATTTAAGCCGG 1620  
DB 24633 GTACGAGGTTAAGAGGGAATCAGGCAAGGCTTCTCAGAGAGGTGACATTTAAGCCGG 24692  
QY 1621 ACATGAAGGATGAACAGGTTAGTTTACCAAGGATGGATGGAAAGGGGTGAGATGATG 1680  
DB 24693 ACATGAAGGATGAACAGGTTAGTTTACCAAGGATGGATGGAAAGGGGTGAGATGATG 24752  
QY 1681 AGGCAAGGGAATCTGAGGATCATAGGCTTAGAGGGGATCTCTGACGCGCTTTGAGGAAG 1740  
DB 24753 AGGCAAGGGAATCTGAGGATCATAGGCTTAGAGGGGATCTCTGACGCGCTTTGAGGAAG 24812



QY 1741 TGAGAGACAGCAGCGAGTCGTAGTGGTTAAAGTAAACAAGCTGAGAGCAGGGAAT 1800  
Db |  
24813 TGAGAGAGACAGCGCAGTCGTAGTGGTTAAAGTAAACAAGCTGAGAGCAGGGAAT 24872  
QY 1801 CCCTGGTCATCAGGGCCTGTGAGTCAGCTCAGAGTGTCTGGGCTTTTGTCTTCTGGG 1860  
Db |  
24873 CCCTGGTCATCAGGGCCTGTGAGTCAGCTCAGAGTGTCTGGGCTTTTGTCTTCTGGG 24932  
QY 1861 AGCAGTCGATTTTAAAGCAGGGAACAGCTGTATTACAGCTGGGAAGATCCTGTGTTGCT 1920  
Db |  
24933 AGCAGTCGATTTTAAAGCAGGGAACAGCTGTATTACAGCTGGGAAGATCCTGTGTTGCT 24992  
QY 1921 GCCTGAAGGGATGAAACTGAGGCTAGGAGCCAGGCTGATAGGAGGATCCAGGGTGA 1980  
Db |  
24993 GCCTGAAGGGATGAAACTGAGGCTAGGAGCCAGGCTGATAGGAGGATCCAGGGTGA 25052  
QY 1981 TGGGAGGCTGGGAGTCCGCGGTGATGACACAGGCTGGGCGCAGGGGATGGGAGGAA 2040  
Db |  
25053 TGGGAGGCTGGGAGTCCGCGGTGATGACACAGGCTGGGCGCAGGGGATGGGAGGAA 25112  
QY 2041 GGAGTAATTGGAGAGGCTGGGCTCTGGCGGAGGAATGATGGTGGCTGAAACAGGG 2100  
Db |  
25113 GGAGTAATTGGAGAGGCTGGGCTCTGGCGGAGGAATGATGGTGGCTGAAACAGGG 25172  
QY 2101 AGAGGAGATGCTTAGGCCACTTTGGAAACAGTAGGCGCAAGGACAGGACACCCCAAG 2160  
Db |  
25173 AGAGGAGATGCTTAGGCCACTTTGGAAACAGTAGGCGCAAGGACAGGACACCCCAAG 25232  
QY 2161 GGGAGGTGCCAAGAGACCAAGCAGAGCTGGCATTTGACAGGGAAGTCTGTCTGGAGCA 2220  
Db |  
25233 GGGAGGTGCCAAGAGACCAAGCAGAGCTGGCATTTGACAGGGAAGTCTGTCTGGAGCA 25292  
QY 2221 GGTCTCTGGTAAGGAGGAAATGGTGAGTTCATCTCCCTCTCTGTCTCAACC 2280  
Db |  
25293 GGTCTCTGGTAAGGAGGAAATGGTGAGTTCATCTCCCTCTCTGTCTCAACC 25352  
QY 2281 TCTAAACTACATGGGCGACAGGACCCAGTGGGATCCCAATATGATGGGATGGTGGATG 2340  
Db |  
25353 TCTAAACTACATGGGCGACAGGACCCAGTGGGATCCCAATATGATGGGATGGTGGATG 25412  
QY 2341 GAAGGAAGGAAGGAGGAAACAACTCTTCATTCCTGGTTATTATACAGAACAGGCCAGG 2400  
Db |  
25413 GAAGGAAGGAAGGAGGAAACAACTCTTCATTCCTGGTTATTATACAGAACAGGCCAGG 25472  
QY 2401 TGGCGTCTCAGCTTGGCATCTAGCACTTTGGAGCTGAGTGGTGGATTAACCTCA 2460  
Db |  
25473 TGGCGTCTCAGCTTGGCATCTAGCACTTTGGAGCTGAGTGGTGGATTAACCTCA 25532  
QY 2461 GGTGAGGATTCAGAGCAGGCTAGACAACTAGACAAACCCCATCTCTACTGAAGATAT 2520  
Db |  
25533 GGTGAGGATTCAGAGCAGGCTAGACAACTAGACAAACCCCATCTCTACTGAAGATAT 25592  
QY 2521 AAAATTAGTGGCGTGTGTCATATGCTGTAAATCCAGCTAGTCGGGAAGCTGAGGCA 2580  
Db |  
25593 AAAATTAGTGGCGTGTGTCATATGCTGTAAATCCAGCTAGTCGGGAAGCTGAGGCA 25652  
QY 2581 GGAGAACTGCTTGAACCCGACAGCAGAGGTTGGGTGAGCTGAGATGTCGCCATTGCAC 2640  
Db |  
25653 GGAGAACTGCTTGAACCCGACAGCAGAGGTTGGGTGAGCTGAGATGTCGCCATTGCAC 25712  
QY 2641 TCCAGCCTGGGTGACAAAGCAAGACCTGCTCAATATAATAATAATAATAATAATAATAA 2700  
Db |  
25713 TCCAGCCTGGGTGACAAAGCAAGACCTGCTCAATATAATAATAATAATAATAATAATAA 25772  
QY 2701 GGAGCCTGGGTGATCCCAAGCTACCTACTTTTTCAGGAGAAATGACTCCCTTACCCCAAGGGC 2760  
Db |  
25773 GGAGCCTGGGTGATCCCAAGCTACCTACTTTTTCAGGAGAAATGACTCCCTTACCCCAAGGGC 25832  
QY 2761 AAAAGATGGGAGAACGATTTGATATGCAATTTATGAGACCTACTGAGTCTCATCCC 2820  
Db |  
25833 AAAAGATGGGAGAACGATTTGATATGCAATTTATGAGACCTACTGAGTCTCATCCC 25892  
QY 2821 TGGGCTAGGCTGGAATGGACTCAGATGGAGCTGAAGAGTCCCCCTCAGGGAACTCACT 2880

Db 25893 TGGGCTAGGCTGGAAATGGAATCAGATGGAGCTGAAGAGTCCCCCTCAGGGAACTCACT 25952  
QY 2881 AGAAGAAGAGGAGGAATCGCGCGGCGGCTGCTCAGCCCTGTAAATCCCAACACTTTGGG 2940  
Db |  
25953 AGAAGAAGAGGAGGAATCGCGCGGCGGCTGCTCAGCCCTGTAAATCCCAACACTTTGGG 26012  
QY 2941 AGGCTGAGGTGGTGGATCAACAAGTCAGGAGATCGAGACCATCTGCTTAACAGTGA 3000  
Db |  
26013 AGGCTGAGGTGGTGGATCAACAAGTCAGGAGATCGAGACCATCTGCTTAACAGTGA 26072  
QY 3001 A 3001  
Db 26073 A 26073

RESULT 2  
AC058798/c  
LOCUS AC058798 173251 bp DNA linear HTG 19-APR-2000  
DEFINITION Homo sapiens clone RP11-166H18, WORKING DRAFT SEQUENCE, 35  
unordered pieces.  
ACCESSION AC058798  
VERSION AC058798.1 GI:7596893  
KEYWORDS HTG; HTGS PHASE1; HTGS\_DRAFT.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 173251)  
AUTHORS Birren, B., Linton, L., Nusbaum, C. and Lander, E.  
TITLE Homo sapiens, clone RP11-166H18  
JOURNAL Unpublished  
AUTHORS 2 (bases 1 to 173251)  
REFERENCE 1 (bases 1 to 173251)  
AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,  
Anderson, S., Baldwin, J., Barna, N., Bastien, V., Bedalov, F.,  
Boguslavskiy, L., Bouckhgalter, B., Brown, A., Burkett, G.,  
Campiano, A., Castle, A., Choepel, Y., Colangelo, M., Collins, S.,  
Collins, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S.,  
Dodge, S., Domino, M., Doyle, M., Ferreira, P., FitzHugh, W., Gage, D.,  
Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,  
Grand-Pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L.,  
Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,  
Klein, J., LaRocque, K., Lamazares, R., Landers, T., Lechoczy, J.,  
Levine, R., Liu, G., Locke, K., Macdonald, P., Marquis, N.,  
McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McPheeters, R.,  
Meldrum, J., Meneus, L., Mihova, T., Miranda, C., Mleaga, V., Morrow, J.,  
Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,  
O'Neill, D., Oliver, T.M., Oliver, J., Peterson, K., Pierre, N.,  
Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,  
Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,  
Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,  
Tefaye, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J.,  
Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J.,  
Young, G., Zainoun, J., Zimmer, A. and Zody, M.  
Direct Submission  
Submitted (19-APR-2000) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)

----- Project Information

Center project name: L7830

Center clone name: 166\_H\_18

----- Summary Statistics

Sequencing vector: M13; M7815; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 156916 bases at least Q40





Db	120951	AGAGAGAGATGCTTAGGCCACTTTGGACACAGTAGGCGACAGGACAGGACACCCCAAG	120892
Qy	2161	GGGAAGTGCCTCCAGAGACACACAGAGCTGGCATTTGGACAGGAAGGCTGTCTGGAGCA	2220
Db	120891	GGGAAGTGCCTCCAGAGACACACAGAGCTGGCATTTGGACAGGAAGGCTGTCTGGAGCA	120832
Qy	2221	GGTGTCTGGATTAAGGAGGAGAAATGTTGACGTTCCATCTCTCTCTCTCTCTCTCAACC	2280
Db	120831	GGTGTCTGGATTAAGGAGGAGAAATGTTGACGTTCCATCTCTCTCTCTCTCTCTCAACC	120772
Qy	2281	TCCTAACTACATGGGSCACAGACACAGGACCTGCGGACCTCCATAATGATGGATGGGTGGATG	2340
Db	120771	TCCTAACTACATGGGSCACAGACACAGGACCTGCGGACCTCCATAATGATGGATGGGTGGATG	120712
Qy	2341	GAAGGAAGGAGGAGGAGAAACAATCTTTCATTCATCTGCTGTTATTTACAGAAAGGCCAGG	2400
Db	120711	GAAGGAAGGAGGAGGAGAAACAATCTTTCATTCATCTGCTGTTATTTACAGAAAGGCCAGG	120652
Qy	2401	TGCGGTGCTCAGCTTGGCTTTCAGCACTTTGGGAGGCTGAGGTGGGTGGATTACTCTCA	2460
Db	120651	TGCGGTGCTCAGCTTGGCTTTCAGCACTTTGGGAGGCTGAGGTGGGTGGATTACTCTCA	120592
Qy	2461	GCTCAGGAGTTCAAGACACAGCTTAGACAAAGCTAGACAAAGCTAGACAAAGCTAGACAA	2520
Db	120591	GCTCAGGAGTTCAAGACACAGCTTAGACAAAGCTAGACAAAGCTAGACAAAGCTAGACAA	120532
Qy	2521	AAAAATAGTGGCGTGGATGAGCATGCTCTGTAATCCAGCTAGTGGGAGCTGAGGCA	2580
Db	120531	AAAAATAGTGGCGTGGATGAGCATGCTCTGTAATCCAGCTAGTGGGAGCTGAGGCA	120472
Qy	2581	GGAAATCCTTTGAACCCAGAGGAGGAGGTTGCGGTGAGCTGAGATCGTGCCTATGCGAC	2640
Db	120471	GGAAATCCTTTGAACCCAGAGGAGGAGGTTGCGGTGAGCTGAGATCGTGCCTATGCGAC	120412
Qy	2641	TCAGGCTGGGTGACAAAGCAAGACCTCGTCTCAATAATAATAATAATAATAATAACAGAA	2700
Db	120411	TCAGGCTGGGTGACAAAGCAAGACCTCGTCTCAATAATAATAATAATAATAATAACAGAA	120352
Qy	2701	GGAGCTGGGTCTATCCAGCTACCTACTTTTTCAGGAGAAATGTTACTCTCTTACCAAGGCG	2760
Db	120351	GGAGCTGGGTCTATCCAGCTACCTACTTTTTCAGGAGAAATGTTACTCTCTTACCAAGGCG	120292
Qy	2761	AAAGGATGGGAGGAGGAGGAGGTTGATGATGATGATGATGATGATGATGATGATGATGATG	2820
Db	120291	AAAGGATGGGAGGAGGAGGAGGTTGATGATGATGATGATGATGATGATGATGATGATGATG	120232
Qy	2821	TGGGCTAGGCTGAATGGAGCTCAGATGGAGCTGAGAGTCCCGCTCAGGGAACCTCACT	2880
Db	120231	TGGGCTAGGCTGAATGGAGCTCAGATGGAGCTGAGAGTCCCGCTCAGGGAACCTCACT	120172
Qy	2881	AGAAAGAGGAGGAAATCGGCGGCGGCGGTGGCTCAGCCTGTATCCCAACACTTTGGG	2940
Db	120171	AGAAAGAGGAGGAAATCGGCGGCGGCGGTGGCTCAGCCTGTATCCCAACACTTTGGG	120112
Qy	2941	AGCTGAGGTGGGTGGATTCACAGGTTCAGAGATCGAGACCATCTCGGTAAACACAGTGA	3000
Db	120111	AGCTGAGGTGGGTGGATTCACAGGTTCAGAGATCGAGACCATCTCGGTAAACACAGTGA	120052
Qy	3001	A 3001	
Db	120051	A 120051	
RESULT 3			
AL714004/c			
LOCUS			
DEFINITION Homo sapiens clone XX-CHR_20-17119-3, *** SEQUENCING IN PROGRESS			
AL714004			
***, 89 unordered pieces.			
ACCESSION AL714004			
VERSION AL714004.35 GI:21104209			
KEYWORDS HTG; HTGS PHASE1.			
SOURCE Homo sapiens (human)			
ORGANISM Homo sapiens			

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

Plumb.R.

Direct Submission

Submitted (20-MAY-2002) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk

On May 22, 2002 this sequence version replaced gi:21068608.

----- Genome Center

Center: Wellcome Trust Sanger Institute

Center code: SC

Web site: http://www.sanger.ac.uk

Contact: humquery@sanger.ac.uk

----- Project Information

Center project name: Chr 20-17119-3

----- Summary Statistics

Assembly program: XGAP4; version 4.5

Sequencing vector: plasmid; L08752; 100% of reads

Chemistry: Dye-terminator; 100% of reads

Consensus quality: 254683 bases at least Q40

Consensus quality: 281730 bases at least Q30

Consensus quality: 299177 bases at least Q20

Insert size: 312719; sum-of-contigs

Quality coverage: 1.63x in Q20 bases; sum-of-contigs

-----

\* NOTE: This is a 'working draft' sequence. It currently consists of 89 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

\* 1 4746: contig of 4746 bp in length

\* 4846: gap of 100 bp

\* 4847 7172: contig of 2326 bp in length

\* 7173 7272: gap of 100 bp

\* 7273 9289: contig of 2017 bp in length

\* 9290 9389: gap of 100 bp

\* 9390 11807: contig of 2418 bp in length

\* 11808 11907: gap of 100 bp

\* 11908 13962: contig of 2055 bp in length

\* 13963 14062: gap of 100 bp

\* 14063 16347: contig of 2285 bp in length

\* 16348 16448: gap of 100 bp

\* 16448 18532: contig of 2085 bp in length

\* 18533 18632: gap of 100 bp

\* 18633 20654: contig of 2022 bp in length

\* 20655 20754: gap of 100 bp

\* 20755 23021: contig of 2267 bp in length

\* 23022 23121: gap of 100 bp

\* 23122 25134: contig of 2013 bp in length

\* 25135 25234: gap of 100 bp

\* 25235 28285: contig of 3051 bp in length

\* 28286 30542: gap of 100 bp

\* 30543 30642: gap of 100 bp

\* 30643 33433: contig of 2791 bp in length

\* 33434 33533: gap of 100 bp

\* 33534 35599: contig of 2066 bp in length

\* 35600 35699: gap of 100 bp

\* 35700 39866: contig of 4167 bp in length

\* 39867 39966: gap of 100 bp

\* 39967 50799: contig of 10833 bp in length

\* 50800 50899: gap of 100 bp

\* 50900 55372: contig of 4473 bp in length

\* 55373 55472: gap of 100 bp

\* 55473 58398: contig of 2926 bp in length

\* 58399 58498: gap of 100 bp

\* 58499 61834: contig of 3336 bp in length

\* 61835 61934: gap of 100 bp

\* 61935 72530: contig of 10596 bp in length

```
* 72531 72630: gap of 100 bp
* 72631 77695: contig of 5065 bp in length
* 77695 77795: gap of 100 bp
* 77795 80590: contig of 2795 bp in length
* 80590 80591: gap of 100 bp
* 80591 80592: contig of 4362 bp in length
* 80592 85152: gap of 100 bp
* 85152 85153: contig of 4135 bp in length
* 85153 89287: contig of 4135 bp in length
* 89287 91654: contig of 2267 bp in length
* 91654 91754: gap of 100 bp
* 91754 94398: contig of 2644 bp in length
* 94398 94499: gap of 100 bp
* 94499 96771: contig of 2273 bp in length
* 96771 96872: gap of 100 bp
* 96872 103634: contig of 6763 bp in length
* 103634 103734: gap of 100 bp
* 103734 107935: contig of 4201 bp in length
* 107935 108035: gap of 100 bp
* 108035 113903: contig of 5868 bp in length
* 113903 114003: gap of 100 bp
* 114003 116208: contig of 2205 bp in length
* 116208 116309: gap of 100 bp
* 116309 118998: contig of 2690 bp in length
* 118998 122033: contig of 2935 bp in length
* 122033 122133: gap of 100 bp
* 122133 124238: contig of 2105 bp in length
* 124238 124339: gap of 100 bp
* 124339 126514: contig of 2176 bp in length
* 126514 126615: gap of 100 bp
* 126615 136561: contig of 9947 bp in length
* 136561 139137: contig of 2476 bp in length
* 139137 139237: gap of 100 bp
* 139237 142728: contig of 3491 bp in length
* 142728 142829: gap of 100 bp
* 142829 165779: contig of 22951 bp in length
* 165779 165879: gap of 100 bp
* 165879 168731: contig of 2852 bp in length
* 168731 168832: gap of 100 bp
* 168832 170924: contig of 2093 bp in length
* 170924 171024: gap of 100 bp
* 171024 173244: contig of 2220 bp in length
* 173244 173345: gap of 100 bp
* 173345 176010: contig of 2666 bp in length
* 176010 176110: gap of 100 bp
* 176110 178420: contig of 2310 bp in length
* 178420 178520: gap of 100 bp
* 178520 181160: contig of 2640 bp in length
* 181160 181260: gap of 100 bp
* 181260 184022: contig of 2762 bp in length
* 184022 184122: gap of 100 bp
* 184122 186233: contig of 2111 bp in length
* 186233 186333: gap of 100 bp
* 186333 188474: contig of 2141 bp in length
* 188474 188574: gap of 100 bp
* 188574 190834: contig of 2260 bp in length
* 190834 190934: gap of 100 bp
* 190934 193769: contig of 2835 bp in length
* 193769 193869: gap of 100 bp
* 193869 196119: contig of 2250 bp in length
* 196119 196219: gap of 100 bp
* 196219 200732: contig of 4513 bp in length
* 200732 200832: gap of 100 bp
* 200832 204093: contig of 3261 bp in length
* 204093 204193: gap of 100 bp
* 204193 206520: contig of 2327 bp in length
* 206520 206620: gap of 100 bp
* 206620 208679: contig of 2059 bp in length
* 208679 208779: gap of 100 bp
* 208779 210932: contig of 2153 bp in length
* 210932 211032: gap of 100 bp
```

```
* 211033 213133: contig of 2101 bp in length
* 213133 213233: gap of 100 bp
* 213233 217420: contig of 4187 bp in length
* 217420 217520: gap of 100 bp
* 217520 222145: contig of 4625 bp in length
* 222145 222146: gap of 100 bp
* 222146 224782: contig of 2537 bp in length
* 224782 224882: gap of 100 bp
* 224882 227854: contig of 2972 bp in length
* 227854 227954: gap of 100 bp
* 227954 230320: contig of 2366 bp in length
* 230320 230421: gap of 100 bp
* 230421 232905: contig of 2485 bp in length
* 232905 233005: gap of 100 bp
* 233005 236533: contig of 3528 bp in length
* 236533 236633: gap of 100 bp
* 236633 239252: contig of 2619 bp in length
* 239252 239353: gap of 100 bp
* 239353 248789: contig of 9437 bp in length
* 248789 248890: gap of 100 bp
* 248890 250940: contig of 2051 bp in length
* 250940 251041: gap of 100 bp
* 251041 253162: contig of 2122 bp in length
* 253162 253262: gap of 100 bp
* 253262 255411: contig of 2149 bp in length
* 255411 255511: gap of 100 bp
* 255511 258118: contig of 2607 bp in length
* 258118 258219: gap of 100 bp
* 258219 260752: contig of 2533 bp in length
* 260752 260852: gap of 100 bp
* 260852 263142: contig of 2291 bp in length
* 263142 263242: gap of 100 bp
* 263242 265368: contig of 2126 bp in length
* 265368 268121: contig of 2653 bp in length
* 268121 268221: gap of 100 bp
* 268221 270223: contig of 2002 bp in length
* 270223 270323: gap of 100 bp
* 270323 274441: contig of 4118 bp in length
* 274441 274541: gap of 100 bp
* 274541 276791: contig of 2250 bp in length
* 276791 276891: gap of 100 bp
* 276891 280982: contig of 4091 bp in length
* 280982 281082: gap of 100 bp
* 281082 283442: contig of 2360 bp in length
* 283442 283542: gap of 100 bp
* 283542 283649: gap of 100 bp

Query Match
Best Local Similarity 16.0%; Score 479.8; DB 2; Length 321519;
Matches 553; Conservative 0; Mismatches 37; Indels 11; Gaps 4;

QY 2408 CTCACGCTGCCATTCTAGCACTTTGGAGGCTGAGGTGGGTGATTAACCTCAGGTGAGG 2467
Db 238884 CTCACGCTGTAATCCAGCACTTTGGAGGCTGAGGTGGGTGATTAACCTCAGGTGAGG 238827
QY 2468 AGTTCAGACAGCCTAGACACAGTACAGAAACCCCATCTCTACTAGATATAAATA 2527
Db 238826 AGTTCAGACAGCCTAGACACAGTACAGAAACCCCATCTCTACTAGATATAAATA 238769
QY 2528 GCTGGCGGTAGTGGCATATGCTGTGTAATCCAGCTAGTGGGAGCTGAGGAGGAGGAGAT 2587
Db 238768 GCTGGCGGTAGTGGCGGCTTCTGTGTAATCCAGCTAGTGGGAGCTGAGGAGGAGAT 238709
QY 2588 CGCTTGA-----ACCCGAGAGGAGGAGGTTGGGTGAGCTGAGATCGTCCCATTTGCAC 2641
Db 238708 CGCTTGAACCTGGGACCCGAGAGGAGGAGGTTGGGTGAGCTGAGATCGTCCCATTTGCAC 238649
QY 2642 CAGGCTGGGTGACAAAGCAAGACCTCGTCTCAATAATAATAATAATAATAATAATAATA 2701
Db 238648 CCAGCCTGGGTGACAAAGCAAGACCTCGTCTCAATAATAATAATAATAATAATAATAATA 238589
QY 2702 GAGCCTGGGTGATCCAGCTACTCTCTTTCAGGAGATGTTACTCCCTTACCCAGGCA 2761
```

Db	238588	GAGCTGGGTCATCCAGCTACCTACTTTTCAGGAGAATGTACTCCCTTACCCAAGGCA	238529
Qy	2762	AAGGATGGGAGAACCAAGTTTGATATGCAATTTATTGAGCACTTACTGAGTCCTCATCCCT	2821
Db	238528	AAGGATGGGAGAACCAAGTTTGATATGCAATTTATTGAGCACTTACTGAGTCCTCATCCCT	238469
Qy	2822	GGGCTAGGCTGGAAATGGAATCAGATGGAGCTGGAAGAGTCCCCTCAGGGAACTTCACATA	2881
Db	238468	GGGCTAGGCTGGAAATGGAATCAGATGGAGCTGGAAGAGTCCCCTCAGGGAACTTCACATA	238409
Qy	2882	GAAGAAGGAGGAATCGGCGGGGCGGTGGCTCAGCGCTGTAATCCCAACACTTTGGGA	2941
Db	238408	GAAGAAGGAGGAATCGGCGGGGCGGTGGCTCAGCGCTGTAATCCCAACACTTTGGGA	238349
Qy	2942	GGCTAGGTGGGTGATCAGC--AAGTCAGGAGATCGAGACCACTCCTGGGTAAACACATGCA	3000
Db	238348	GGCTAGGTGGGTGATCAGC--AAGTCAGGAGATCGAGACCACTCCTGGGTAAACACATGCA	238289
Qy	3001	A 3001	
Db	238288	A 238288	

**RESULT 4**

HS999D10	9332 bp	DNA	linear	PRI_04-MAR-2003
LOCUS				
DEFINITION	Human DNA sequence from clone CTA-999D10 on chromosome 22q13.3, complete sequence.			

ACCESSION	Z94802
VERSION	Z94802.1
KEYWORDS	HTG.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
REFERENCE 1 (bases 1 to 9532)

REFERENCE	1 (bases 1 to 9532)
AUTHORS	Clark, G.
TITLE	Direct Submission
JOURNAL	Submitted (04-MAR-2003) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquey@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk On Feb 5, 1998 this sequence version replaced gi:2578153.
COMMENT	

COMMENT

On Feb 5, 1998 this sequence version replaced gi:2578153.  
During sequence assembly data is compared from overlapping clones.  
Where differences are found these are annotated as variations  
together with a note of the overlapping clone name. Note that the  
variation annotation may not be found in the sequence submission  
corresponding to the overlapping clone, as we submit sequences with  
only a small overlap as described above.  
The following abbreviations are used to associate primary accession  
numbers given in the feature table with their source databases:  
Em; EMBL; Sw; SWISSPROT; Tr; TREMBL; Wp; WORMPEP; Information  
http://www.sanger.ac.uk/Projects/C\_elegans/wormpep CTA-999p10 is  
from the human BAC library described in U-J. Kim et al. (1996)  
Genomics 34, 213-218.

Genomics 34, 2137-2140.  
VECTOR: pBelOBAC11  
----- Genome Center  
Center: Wellcome Trust Sanger Institute  
Center code: SC  
Web site: <http://www.sanger.ac.uk>  
Contact: [humquerry@sanger.ac.uk](mailto:humquerry@sanger.ac.uk)

-----

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

This sequence was generated from part of bacterial clone contigs of human chromosome 22, constructed by the Sanger Centre Chromosome 22 Mapping Group. Further information can be found at

<http://www.sanger.ac.uk/HGP/Chr22>.  
 FEATURES  
 source Location/Qualifiers  
 1. .9532  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="22"  
 /map="q13.3"  
 /clone="CTA-999D10"  
 /clone\_lib="CIT978SK-A2"  
 ORIGIN  
 Query Match 9.2%; Score 276.4; DB 9; Length 9532;  
 Best Local Similarity 49.2%; Pred. No. 1.9e-62;  
 Matches 1186; Conservative 0; Mismatches 1161; Indels 63; Gaps 15;  
 314 GCCGGTGCAGTGGCTAGTGCTCGAATCCCAACACATTGGAGCGCGAGGTGGGTGCAAT 373  
 |||||  
 375 GCCAGGCACGGTGGCTCACACTGTATCTTAGCACATTGGAGGCCGAGCGCAGAGGAT 434  
 |||||  
 374 CACTTGAGGTCCAGGAGTTTCGAGACCAGCTCGGCCAAACATGGTGAACCCCTCTCTCTACTA 433  
 |||||  
 435 CACTTGAGTCCAGGAGTTCGAAATCAGCTCGCCAAACATGGTGAACCCAGTCTCTACTA 494  
 |||||  
 434 AAAATATATAAAATTAGCTGGGCATGGTGTGCTACCTGTAATCCAGATACTTTGGGAGG 493  
 |||||  
 495 AACATA-CAAAATTAGCTGGCGCTGGTGGTGGCTGTATCCCACTACTCCGGAGG 553  
 |||||  
 494 TTGAGGCAGGAGATCGCTTCGAACCCGGGAGGCGAGATGTTGCAGTGAACCAAGACTGTGTC 553  
 |||||  
 554 CTGAGGCAGGAGATCACTTGAACCTGGGAGCGGAGGTTTCAGTGAAGTGAAGTCGAGC 613  
 |||||  
 554 CACTGCACTCCAGTCTGGGCCAACAGAGTGAAGCTCCATCTCAAAACAAACAAACAAAGC 613  
 |||||  
 614 CACTGCACTCCAGCTCGGTGACAGCGCAGACTCCGCTCTCA---AAAAAAAAGGCATT 670  
 |||||  
 614 AGTGCCCATCATGTAGAGATTTAGATGATTGATGAGGACTGAGCCTTGTGCAAGTGAACA 673  
 |||||  
 671 TTTTGGAAATAGAAAAACATAATTCTAAAAATTCCTATGGAAATTTTAAGGGAACCCAAAT 730  
 |||||  
 674 CTCACATATCACCGAGTTGTAGTATCAGTGATAACCATCAATGATCCAGGTAAAGCCCTG 733  
 |||||  
 731 GGCCATAACAATCTTTGGAAAAAAAAGTTTACTTCTTGATGACTAAAGATATAAATAGNA 790  
 |||||  
 734 AGGGT--TCAGAAAGATGCCGAGCGCTTCAAGGTGCTGGGGATTTGTTGGCAAGCCCT 791  
 |||||  
 791 AACATAAANAAGAAAGAGATGTAAAGAAACAAATAATTAAGGTTAAAAAAGAACAA 850  
 |||||  
 792 CGAATAATAGAAACAGT-----TCTCTGTATTACACAGAAAGCAGAGGCCCATGCTG 845  
 |||||  
 851 CAAAGACTGGGCACCGTGATTCATGCCGTAAATCCCAAGCATTCTTTGGGAGCCGAGGCGG 910  
 |||||  
 846 GGTGCTGCC--AGGAACTCAGTAGTAACTTAACACAGCACACCGGTGCTGCTTCCCCAGCGCA 903  
 |||||  
 911 GCGGATCAGAGGTACAGAGATTGAGACCATCTCTGGCCAAATGTGTGAACCCCATCTCT 970  
 |||||  
 904 CTTAGGCCAGTGGGGAAAACAGACTCACCACACAGTCCAGCCACAGAGTGGTCAGGGCCAA 963  
 |||||  
 971 ACTAAAAATACAAAAAATTAGCTGGCGGTGGTGGCACGTCCTGTAGTCCAGCTACTCG 103  
 |||||  
 964 GATGGGGAAGCACGGGAGAAAGGTTCAGGGTGGGATGGGAGGGGTTCAGGCCAGAGGGG 102  
 |||||  
 1031 GGAGGCTGAGGCAGGGGAAATTGCTTGAACCCAGGAGCGAGGTTGCACTGAGCCCAAGAT 109  
 |||||  
 1024 TCAGGCCCAAGGCT-----GAGGGAAGCCTGGGACTGTAGGAATTTAGAGAGAGTGA 107  
 |||||  
 1091 TCGGCCCACTGCATTCACAGCTTGGCAACAGCGAGACTCCGTTCGCAAAAAAAGAAAAA 115  
 |||||  
 1075 CCTGACCCCGCATGTTTGGTGAAGGAGATTCAGGAAAGTCTTCTCGAAGAGAGGCTGTG 113  
 |||||  
 1151 AAAAAATTAGCCAGGTGTGGTGTGGACACCTGTATCCCACTACTCGGAGGCGCGAGG 121  
 |||||  
 1135 GAGCTGAGACTCATAGATGAGTGGGAGGGTGTTCAGGCGCAAGACACCGACCTTACA 119  
 |||||



Db 1211 CAGGGAGAAATTCCTTGAATCTGGAGGTGGAGGTTGCACTGAGCTCAGATTATGCCACTG 1270  
QY 1195 AAAGCATGACCTTTGAGAGAGCAATTCATCCATTCACCTGATGAATTTTCAGACTGGGCAC 1254  
Db 1271 CACTCCAGCCTCAACGACAGAGCAAGACTCCCATCTCCAAAAATAAAAAAAGACAGGTGC 1330  
QY 1255 GCTGCTCATGCTGTAATCCAGCACCTTTGGAAGGCTGAAATGGGAGGATGACTTGAGC 1314  
Db 1331 AGTGGCTCAGCCTGTATCCAGGACCTTTGGAGGCTGAGGCCGCGAGATCACCTGAGG 1390  
QY 1315 CTAGGCATTTGTGACAAAGCTGGGCAACATGGTGAGACCCCTGCTCCACAAAAACAAACAA 1374  
Db 1391 TCAGGAGTTCAAGACCAAGCCTGGCCACATGGTGAAACCCCATCTCTACTAAATACAA 1450  
QY 1375 ACAAACAAAAAATCATATATACCTGTAGTACATGGGTACAGGTACATAGAAATGACTCAGG 1434  
Db 1451 AAAATTAGCCTGGGTGATGGGGTGCTGTGTAATCCAGCTACTCTCGGGGGGCTGAGAGC 1510  
QY 1435 CAGATATGGTGT---CCTCTCTCTACTGTGGGAGAGCGGGCTTACTGTCAGTAAGCAA 1491  
Db 1511 GGAGAATTGCTTGAGCCTGGGAGACGGTGTGTCAGTAAGTGTGATGAGATTGCACCATGCT 1570  
QY 1492 TAGAGGAGGGAATATATCTCTAAATGAGAGGTACAGATTGAGAGCAAAACACAGGSCA 1551  
Db 1571 TCCAGCCTGGGACGCTGAGCAAGACTCCATCTCAAAAAAAGAAAAAAGAAAGGCT 1630  
QY 1552 CAGGCATATGTACAGAGGTAAAGAGGNAATCAGGAGGCTTCTCAGAGAGGTGACATT 1611  
Db 1631 GGGTGAGTGGCTCATGCTATATATCCAGCACTTTTGGGAGGCGGAGACGGGTGGATCAT 1690  
QY 1612 TAAGCCGGGACATGAAGGATGAACGAGTTA-----GTTCAACCAAGGATGGATGMAAGG 1666  
Db 1691 GAGTCAAGGAGAGGAGACCATCTGGCTACATGTTGTAACCTGCTCTCTACTATAAAAT 1750  
QY 1667 GGTGAGAGTATGAGGAGCAGAGGAACTGCAAGATCATAGGCCCTPAGACAGGGGCTCTGA 1726  
Db 1751 ACAAAAAATTTAGCCGGCGTGTGGCGGGCGCTATAGTCCAGCTACTCGGGAGGCTGA 1810  
QY 1727 GGCCTTTGAGGAGTGAAGAGACCAAGCAGCTGCTAGTGGTAAAGTAAACAAAGCTGA 1786  
Db 1811 GGC-----AGGAGATGGCTGAACCCGGGAGGAGAGCTTTGAGTGAGCAAGATCGCG 1865  
QY 1787 GAAGCCAGGGAATCCCTGCTCATGCAAGGCGCTGTGAGTCAGCTCAGCTGTTGGGCTT 1846  
Db 1866 CCATTGCACTCCAGCCTGGCCACAGAGCGAACTCCATCTCAAAAAAAGAAAAA 1925  
QY 1847 TTGTTTTCTGGAGCAGTCGATTTTAAAGCGGAAACAGCTGTATTTTCAGAGTTGGGAAG 1906  
Db 1926 AAGAACAAGTTGGAGGACTCTCATTTGTCAATTTTAAATACTTACCACAAGCTACATA 1985  
QY 1907 ATCCTGTGGTTGCTGCTGAAGGGATGAACCTGGAGGCTAGAGGCCAGGGGTGATAGG 1966  
Db 1986 AT-----CAAAACAGCATGGCGAGCATGGTGGCTCACACCTGTAATCCCA 2031  
QY 1967 AGGATCAGGGTATGGGAGGCTGGGAGGTCGCGGTGATGACACAGGCTGGGGCCAG 2026  
Db 2032 GCACCTTTGGAGGCTGAGGAGGTGATCCTGAGT-----CAGGAGTTCAAGACC 2084  
QY 2027 GGGATGGGAGGAGGATTAATTTGGAGAGGCTGGGGCTCTGGCCGAGGAATGGATGGT 2086  
Db 2085 AGCCTGGCCAAACATGTTGAACTCCATCTCTACTAAAAATACAAAAATAGCTGGGTG 2144  
QY 2087 GGGCTGAAACAGGAGGAGGAGATGCTTAGGCCACTTTGG--AACAGTAGGGCAAGGA 2145  
Db 2145 GTGGCAAAAGCCTGTATCCAGCTACTCGGGAGGCTGAGGAGGAGAAATGCTTGAGCC 2204  
QY 2146 CAGGAGACACCAAGGGGAAGTCCCAAGAGACACAGACAGGCTGCAATTTGGACAGGAA 2205  
Db 2205 CGGAGGACAGGTTGCACTGAGTGCATGAGTACGACACTGCACTCCAGCTGACTGACAG 2264  
QY 2206 GGTCTCTGGAGCAGGTGTCTTTGGATAAGGGAGGAAATGGTGCAAGTTCCATCCCTCCTC 2265

Db 2265 AGTGAGACTGTCTCAAAAACAAAAACAAAAACAAAAACAGTGTGGCACTGGCTTTAAAGACAG 2324  
QY 2266 CTTCTCTGTTCAACCTCTAAACTACATGGGGCACAGGACCCAGTGGAGTCCATTAATCA 2325  
Db 2325 ACACAGAGATCAAT---GGGGTTAAATAGACAGCCAGCAAAATAAACACTTCAATATATGG 2381  
QY 2326 TGGGATGGTGTATGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2385  
Db 2382 TCAATGGTTTTCAACAAGGGTGCCAGAAATCATTCATGGTGGAAAGAAATGATCTTTTCA 2441  
QY 2386 ACAGAAACAGGCGAGCTGGCTGGT---GCTCAGCTTGCATTTAGCACTTTGGAGGCTGAGG 2444  
Db 2442 ACAATGGGGCAGAGTGGTGGCTCAGCCTGTAAATCCTAGCACTCTGGGAGGCGGAGG 2501  
QY 2445 TGGGTGGATTACCTCAGGTCAGGAGTTCAAGACCAAGCCTAGACAACTGAGAGAAACCCCA 2504  
Db 2502 TGGGCACATCATCTGAGTCAAGGAGTTTAAAGACCAAGCTGCGCAACGTTGGTGAATCTCTG 2561  
QY 2505 TCTCTACTGAAGATATAAA---ATTAGCTGGGGTGTAGTGCAATGCTCTTAATCCAGCTA 2563  
Db 2562 TCTTGACTGAAAAATACAAACATTTAGCTGGGCATGGTGGTGCATGCTGTAAATTCAGCTA 2621  
QY 2564 GTCCGGAGCTGAGCGAGGAGAAATCGCTTGAACCCGAGGAGGAGGTTGCGGTGAGCTG 2623  
Db 2622 CTGAGAGGCTGAGCGAGGAGAAATCATCTGAGCCCGGAGGAGGTTGCAAGTGAGCGG 2681  
QY 2624 AGATCGTGCCATTGCACTCCAGCCTGGGTGCAAAAGCAAGACCTCTCTCAATTAATA 2683  
Db 2682 AGATCTCGCACTGCACTCAAGCTTGGACACAGACAGGAGGAGGAGGAGGAGGAGGAGG 2741  
QY 2684 ATATTACAA 2693  
Db 2742 CAAGCAAAA 2751

## RESULT 5

AL359916

LOCUS  
DEFINITION

AL359916 126053 bp DNA linear PRI 12-FEB-2001  
Human DNA sequence from clone Rpl1-55008 on chromosome 20. Contains  
a novel gene encoding a protein kinase, an RPL7 (60S Ribosomal  
Protein L7) pseudogene, a CpG island, ESTs, STSs and GSSs, complete  
sequence.

## ACCESSION

AL359916

VERSION

AL359916.8

KEYWORDS

HTG; CpG island; protein kinase; RPL7.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 126053)

Griffiths, C.

Direct Submission

Submitted (08-FEB-2001)

Sanger Centre, Hinxton, Cambridgeshire,

CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk

requests: clonerequest@sanger.ac.uk

On Aug 22, 2000 this sequence version replaced gi:9801431.

During sequence assembly data is compared from overlapping clones.

Where differences are found these are annotated as variations

together with a note of the overlapping clone name. Note that the

variation annotation may not be found in the sequence submission

corresponding to the overlapping clone, as we submit sequences with

only a small overlap as described above.

The following abbreviations are used to associate primary accession

numbers given in the feature table with their source databases:

Em: EMBL; Sw: SWISSPROT; Tr: TrEMBL; Wp: WORMPEP; Information

on the WORMPEP database can be found at

http://www.sanger.ac.uk/Projects/C/elegans/wormpep This sequence

was generated from part of bacterial clone contigs of human

chromosome 20, constructed by the Sanger Centre Chromosome 20

Mapping Group. Further information can be found at

http://www.sanger.ac.uk/HGP/Chr20

IMPORTANT: This sequence is not the entire insert of clone

RPl1-55008 It may be shorter because we sequence overlapping

sections only once, except for a 100 base overlap. The true left end of clone RP11-128M1 is at 125354 in this sequence. The true right end of clone RP4-68402 is at 100 in this sequence. This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. RP11-55008 is from the library RPCL1-11.2 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>  
VECTOR: pBACE3.6.

## FEATURES

source

```

1. 126053
   /organism="Homo sapiens"
   /mol_type="genomic DNA"
   /db_xref="taxon:9606"
   /chromosome="20"
   /clone="RP11-550O8"
   /clone_lib="RPCI-11"
3. 106

```

repeat_region	3. .106
misc_feature	/note="MIR repeat: matches 51. .154 of consensus" complement(815. .1340)
repeat_region	/note="match: GSS: Em:AQ672921" 971. .1046
repeat_region	/note="MER5B repeat: matches 52. .131 of consensus" 1095. .1222
misc_feature	/note="L2 repeat: matches 2407. .2551 of consensus" 1341. .1847
repeat_region	/note="match: GSS: Em:AQ457583" 1749. .1819
repeat_region	/note="L1MA7 repeat: matches 6214. .6287 of consensus" 1930. .2023
repeat_region	/note="MER2 repeat: matches 110. .206 of consensus" 2024. .3568
repeat_region	/note="MER52A repeat: matches 1. .1755 of consensus" 3567. .3586
repeat_region	/note="MER2 repeat: matches 96. .115 of consensus" 3587. .3914
repeat_region	/note="MER2 repeat: matches 2. .344 of consensus" 3915. .4010
repeat_region	/note="MER2 repeat: matches 1. .96 of consensus" 4011. .4603
misc_feature	/note="L1MA8 repeat: matches 5632. .6221 of consensus" complement(4480. .4941)
repeat_region	/note="match: GSS: Em:AQ632384" 5309. .5412
repeat_region	/note="MLR1G repeat: matches 412. .512 of consensus" 5608. .5867
repeat_region	/note="MLR1G repeat: matches 11. .271 of consensus" 6881. .7193
repeat_region	/note="AluY repeat: matches 1. .303 of consensus" 7936. .8016
repeat_region	/note="MIR repeat: matches 87. .167 of consensus" 8241. .8270
repeat_region	/note="15 copies 2 mer ca 100% conserved" 9748. .10266
repeat_region	/note="L1MC4 repeat: matches 7176. .7719 of consensus" 10267. .10376
repeat_region	/note="L1MC4 repeat: matches 7866. .7977 of consensus" 10426. .10566
repeat_region	/note="MIR repeat: matches 91. .236 of consensus" 11561. .11783
repeat_region	/note="MIR repeat: matches 21. .262 of consensus" 12954. .13183
repeat_region	/note="MIR repeat: matches 7. .262 of consensus" 13236. .13286
repeat_region	/note="L2 repeat: matches 2445. .2495 of consensus" 13345. .13421
repeat_region	/note="L2 repeat: matches 2232. .2288 of consensus" 13443. .13575

```
repeat_region 29289..29369
/Note="MIR repeat: matches 73. .152 of consensus"
repeat_region 29387..29627
/Note="MIR repeat: matches 8. .259 of consensus"
repeat_region 29732..30251
/Note="MLTIG repeat: matches 12. .509 of consensus"
repeat_region 30586..30685
/Note="MIR repeat: matches 90. .192 of consensus"
repeat_region 30934..31026
/Note="L1M4 repeat: matches 7864. .7949 of consensus"
repeat_region 31072..31167
/Note="48 copies 2 mer ac 69% conserved"
repeat_region 31188..31357
/Note="MER20 repeat: matches 38. .218 of consensus"
repeat_region 32210..32504
/Note="AluY repeat: matches 1. .291 of consensus"
repeat_region 32506..32559
/Note="I2 repeat: matches 2645. .2697 of consensus"

Query Match 9.1%; Score 272.2; DB 9; Length 126053;
Best Local Similarity 48.0%; Pred. No. 2.4e-61;
Matches 1298; Conservative 0; Mismatches 1353; Indels 54; Gaps 16;

QY 332 TGCCTGAATCCCAACCTTTGGAGCGGAGTGGGTGGATCACTTGGAGTCAGGAGTT 391
DB 46550 TGCCTGAATCCCAACCTTTGGAGCGTGGGTGGCGGATCA--TGAGTCAAGGTT 46607

QY 392 CGAGCAGAGCTGCCACACCTGGTGAACCTGTCTTACTTAAATAATATAATAGCT 451
DB 46608 CGAGCAGAGCTGCCACACCTGGTGAACCTGTCTTACTTAAATAATATAATAGCT 46667

QY 452 GGGCATGGTGGTACCTGTAACTCCAGATCTTGGGAGTTGAGCGAGGAGTCCG 511
DB 46668 AGGCATGGTGGGTCCTGTAACT--CGAGTCTCGGAGGCTGAGCGAGGAGTCC 46726

QY 512 TTGAACCCGGAGGAGAGTGTTCAGTGAACCAAGACTGTGCCACTGCATCTCAGTCTGG 571
DB 46727 TTGAACCCGGAGGAGAGTGTTCAGTGAACCAAGACTGTGCCACTGCATCTCAGTCTGG 46786

QY 572 GCAACAGAGTGCCCTCCATCTCAACCAACCAACCAACCAACCAACCAACCAACCAAC 630
DB 46787 GTGACAGAGTGAGCTCCATCTCAACCAACCAACCAACCAACCAACCAACCAACCAAC 46846

QY 631 -----ATTGAGTGAATGAGTGAAGAGTGAAGCTGTGCAAGAGTGAAGTCAATCA 684
DB 46847 TCATAAATTTGTACAAAGATCCCTTTAAATATAAGACATATAGATTAAGTGA 46906

QY 685 CAGGTTGTAGTATCAGTGAATACCAATCAATGATCAGGTAAAGCCCTGAGGTTTCAGAA 744
DB 46907 AGGGGTAGAGAACATATATGCAATGCAACCAACCAATCAAAAGAAAGCTGAAGCATCTATC 46966

QY 745 AGATGCGGAGGCTTTCAAGGTGCTGGGATTTGGGCAAGCCCTCGATTAATAGAAA 804
DB 46967 TAAATTTCAATAGAGCGCGGCGAGTGGCTACGCTGTAAATCCCGAGCACTTTGGGAGG 47026

QY 805 CAGTCTCTGTATTACACAG--AAAGCAGGAGGCCCATGCTGGGTGCTGCCAGGAAT 861
DB 47027 CTGAGCGGCTGGATTACAGGTCAAGATCAAGACCATCTCTGGCCAAATGGTGAAC 47086

QY 862 CAGTAGTAATAGACAGCACCGGTGCTTCCAGCGCATCTAGGCGCAGTGGGGA 921
DB 47087 CCTGTCTTACTAAGAAATATAAATTTAGTGGCATGGTGAACGCTGTAGTCCCA 47146

QY 922 CAGACTCACCACACTCCAGCCAGAGTGTGTCAGGCGCAAGATGGGGAAGCACGGGA 981
DB 47147 TCTACTCTGGAGCTAGAGCA--AGAGAAATCAATTTGAAATCTGGCAGCGAGGTTGCA 47204

QY 982 GAAAGTTCAGGTTGGATGGGAGGGGTTCAGGSCAAGAGGGGTTCAGGCGCAGGCTCAGGG 1041
DB 47205 GAACTGAGATTGTGCACCTGCACTTAGCTGAGCGACAGATGAAGCTCCGCTCATAA 47264

QY 1042 AAGCCCTGGGACTGTAG--GAATTTAGAGGAGGTACCTGACCCGCGCATGTTTGGTGAAGCA 1100
```

```
DB 47265 CAACAAACAAAGCAGAGCAGACTTTCAGAGCAAGAAAAATTTATCAGAAAAAGAGAGC 47324
QY 1101 GATTTCAGGAAGTCTTCTTCGGAAGAGAGGCTCTCGGAGCTGAGACTCATAAGATGATGGG 1160
DB 47325 ATTACATAACCAATAAATGGTCAATTTCTCCAAAGATACACATAAATAATATGCTCCAG 47384
QY 1161 GAGGGTGTTCAGGCGAGAAAGACCA--GCACCTACAAAAGCATGCTTTGAGAGAAAGCAT 1219
DB 47385 AAGATATAACCAATCTTTAATGCTCTATGCACTTAAACACACAGCATCAAAATACATGAGC 47444
QY 1220 CATCCATTCAACTGATGAATTTTCAGAGCTGGGCGAGCTGGCTCATGCTGTAAATCCAGC 1279
DB 47445 AAAAATCTGACAGAAATGCAAGGAGCGGACAGTGGCTCATGCTGTCAATCCAGC 47504
QY 1280 ACTTTGAAAGCTGAATGGGAGGATGACTTTGAGCCTAGGCAATTTGTGACAAAGCTGGGC 1339
DB 47505 ACTTTGGAGGCGCAAGACAGGTGATCACTTGTGCTCCAGGAGTTTACAGCAGCTGGGA 47564
QY 1340 AACATGCTGAGACCTGCTCCACAAACAAACAAACAA-----CAAAAA 1385
DB 47565 AGTGTAACAAATCCCATCTCTACAAAAATATAAAAAATTTAGCCAGCATGTTGGCTACA 47624
QY 1386 ATCAATTAATCTGTACCATGGGTACAGGTACATAGAAATGACTCAGGCGAGATATGGT 1445
DB 47625 CTGTGTCCAGCTCTCAGAGGCTGACGAGGAGGATCACTTGAGCCCCAGAGTTG 47684
QY 1446 TCCTCTCTACTGTGGAGAGGCGGCTTATATGCAAGTAAAGCAATAGAGGAGGGAAT 1505
DB 47685 AGGCTGAGTGAGTTGTGTTGTGCTGCACTGCACTCAGCTAGGTGACAGAGAGACCT 47744
QY 1506 ATAACTCTAAATGAGAGGTACAGATTTGAGAGCAACACAGGCGACAGGCATATGATG 1565
DB 47745 GTCTCAAAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 47804
QY 1566 AGGTAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1624
DB 47805 CTACTATAGTAGAGTCTCAATCTCTCTATCAGCAGTGGCAAAATCAGCAGGAGGAGG 47864
QY 1625 GAAGGATGAACGAGTTAGTTTCAACAGGATGGGATGGGAGGAGGAGGAGGAGGAGGAGG 1684
DB 47865 AGTTCAGTAGGATATAGTTGAACTCAATAGCACCATTATCAACAGAAATTAATTAACA 47924
QY 1685 AGAGGAGTTCAGGATCATAGGCTTAGAGAGGAGGATCTGACGCGCTTTCAGGAGAGTGA 1744
DB 47925 AATACAGACTACTTCAATCAACCAACCAATAGATACCCATCTCTCAATTTCTGATGAAT 47984
QY 1745 AGAAGACCGAGCTGCTAGTGGGTTAAGTAAACAAAGCTGAGAGCGGAGGAGGAGGAGG 1804
DB 47985 ATCACTATGATAGACTACATTTCTGCTCATTAACACACCTTAAACACATTTAAAGAA 48044
QY 1805 G--GTCAATGAGGCGCTGTGAGTCACTGAGTGTGTTGGGCTTTTGGGCTTTTCTCTGG 1862
DB 48045 GAAGTTGGCGAGCATGTTGGCTCAGCGCTGTAATCCAGCACTTTGTGAGGCGAGGCA 48104
QY 1863 CAGTCAATTTAAGCAGGAGCAAGCTGTATTCAGATTGGGAGATCCTGTGTTGCTGCTG 1922
DB 48105 GCAATATCACTTGTGAGTCAAGGATTCGAGACCGCTGACCAACATGAGGAGGAGGAGGAG 48164
QY 1923 CTGAGGAGGATGAATGAGGAGGCTAGGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1982
DB 48165 TCTACTAAATATCAAAATAGCTGGGATGTTGGGATGTTGGGATGTTGGGATGTTGGGAT 48224
QY 1983 GGGAGGCTGGGAGGCTCGCGGTGATGGACAGGAGGCTGGGCGGAGGAGGAGGAGGAGGAGG 2042
DB 48225 GGGAGGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 48281
QY 2043 AGTAATTTGGAGAGGCTGGGCTCTGGCCGAGGATGATGTTGGGCTGAAACAGGAGGAG 2102
DB 48282 AGATCAGCGCACTGATGCGAGGCTGGGCAACAGAGGAGGAGGAGGAGGAGGAGGAGGAG 48341
QY 2103 AGGAGAGATGCTTAGGCGCACTTTGGAACACAGTAGGCGAGGAGGAGGAGGAGGAGGAG 2162
DB 48342 AAGAAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 48401
```



\* 119646 147254: contig of 27609 bp in length  
 \* 147255 147354: gap of unknown length  
 \* 147355 179607: contig of 32253 bp in length.

## FEATURES

Location/Qualifiers  
 1..179607  
 /organism="Homo sapiens"  
 /db\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="19"  
 /clone="RP11-486P13"  
 misc\_feature  
 1..1338  
 /note="assembly\_name:Contig8"  
 1439..3330  
 /note="assembly\_name:Contig9"  
 3431..5429  
 /note="assembly\_name:Contig10"  
 5530..8107  
 /note="assembly\_name:Contig11"  
 8208..11180  
 /note="assembly\_name:Contig12"  
 11281..14039  
 /note="assembly\_name:Contig13"  
 14140..17230  
 /note="assembly\_name:Contig14"  
 17331..21334  
 /note="assembly\_name:Contig15"  
 clone\_end:SP6  
 vector\_side:left  
 21435..26877  
 /note="assembly\_name:Contig16"  
 26978..31453  
 /note="assembly\_name:Contig17"  
 31554..36515  
 /note="assembly\_name:Contig18"  
 36616..46678  
 /note="assembly\_name:Contig19"  
 46779..57999  
 /note="assembly\_name:Contig20"  
 58100..74742  
 /note="assembly\_name:Contig21"  
 74843..92504  
 /note="assembly\_name:Contig22"  
 92605..119545  
 /note="assembly\_name:Contig23"  
 119646..147254  
 /note="assembly\_name:Contig24"  
 147355..179607  
 /note="assembly\_name:Contig25"

## ORIGIN

Query Match 8.8%; Score 263.8; DB 2; Length 179607;  
 Best Local Similarity 67.8%; Pred. No. 4.2e-59;  
 Matches 414; Conservative 0; Mismatches 192; Indels 5; Gaps 3;  
 QY 2393 AGGCAGGTGGGTCTACGCTTCCATTCTAGCATTCTGGAGGCTGAGGTGGGTGGA 2452  
 DB 25988 AGCCGGGGCGGTCTACGCTGTAATCTTAGCAATTTGGAGTGGAGTGGCGGA 26047  
 QY 2453 TTACTTCAGGTGAGGTTCAAGACCGCTAGAACCTAGAGAACCCCATCTCTACT 2512  
 DB 26048 TCACCTGAGTCAAGGTTTGAAGACCGCTGCGCAACATGCGCAACCCCATCTCTACT 26107  
 QY 2513 GAAGATATAAA-ATTAGCTGGGGTGTAGTCATATCCCTGTATCCAGCTAGTCGGGA 2571  
 DB 26108 AAAATACAAAGATTAGCCGGGTTGGTGGCATGCTCTGTATCTCAGCTACTCGGGAG 26167  
 QY 2572 GCTGAGGAGAGAGATTCGCTTGAACCCGAGAGGAGGTTCGGTGAAGCTCAGATCGTG 2631  
 DB 26168 GCTGAGGAGAGAGATTCGCTTGAACCCGAGAGGAGGTTCGATGAGCCGAGATCATG 26227  
 QY 2632 CCATTCAGCTCAGGCTGGGTGACAAAGCAAGACCTCGTCTCAATA-ATAATAATAA 2690  
 DB 26228 CCATTCAGCTCAGGCTGGGCAACAGAGGAGGAGCTCTACTCAAAATAATAATAA 26287

QY 2691 CAAAACAGAGAGCGCTGGGTCAATCCAGCTACCTACTTTTTCAGGAGATGACTCCCTT 2750  
 DB 26288 AATAAATAAATAAATCATGGGAAGCGAAAAATACATTTATGTGTTTTCCTTCCCTC 26347  
 QY 2751 ACCCAAGGCGCAAGGATGGGAGAACCAAGTTTGAATATGCAATTTATTGAGCACCTACTGAG 2810  
 DB 26348 CCGCAATTAAACCCCTTTTCATTTTCCCATGTTCTCTCGCAGTTCCCATCTCATGACAA 26407  
 QY 2811 TCCTCATCCCTGGGCTAGGCTGGAAATGGACTCAGATGGAGCTGAAGAGTCCCTCAGG 2870  
 DB 26408 ATCTTTGATTCCAGAGCGATAAATAAATATGTCATCATGAGCATCCATCGGCAGACGTG 26467  
 QY 2871 GAACCTCACTAGAAAGAGGAGGAATCGCGCGGCGGTGGCTCACGCTCTGTAATCCCA 2930  
 DB 26468 GCACAAAGTCAATAAACAAAC---TCTAGGCTGGGCACAGTGGCTCACTCTGTAATCCCA 26524  
 QY 2931 ACCTTTGGGAGGCTGAGGTGGGTGGATCAACAGGTGAGGAGATCGAGACCATCTGGCT 2990  
 DB 26525 GCATTTAGGAGCGCGAGGAGGTGCATCAAGGTCAAGAGATCAAGACCATCTGGCC 26584  
 QY 2991 AACACAGTGAA 3001  
 DB 26585 AACATGGTGA 26595

## RESULT 7

HS268H5

LOCUS

DEFINITION

Human DNA sequence from clone CTA-268H5 on chromosome 22q13.2-13.3,

complete sequence.

ACCESSION

AL008718

VERSION

AL008718.23

KEYWORDS

HTG.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

HS268H5 189833 bp DNA linear PRI 04-MAR-2003  
 Human DNA sequence from clone CTA-268H5 on chromosome 22q13.2-13.3,  
 complete sequence.  
 AL008718  
 VERSION AL008718.23 GI:6006479  
 KEYWORDS HTG.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 189833)  
 Direct Submission  
 Submitted (04-MAR-2003) Wellcome Trust Sanger Institute, Hinxton,  
 Cambridgeshire, CB10 1SA, UK. E-mail enquiries:  
 humquery@sanger.ac.uk  
 On Oct 1, 1999 this sequence version replaced gi:6002134.  
 ----- Genome Center  
 Center: Wellcome Trust Sanger Institute  
 Center code: SC  
 Web site: http://www.sanger.ac.uk  
 Contact: humquery@sanger.ac.uk

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em; ENBL; Sw; SWISSPROT; Tr; TREMBL; Wp; WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C\_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 22, constructed by the Sanger Centre Chromosome 22 Mapping Group. Further information can be found at

http://www.sanger.ac.uk/HGP/Chr22	
CTA-268H5 is from the human BAC library described in U-J. Kim et	
al. (1996) Genomics 34, 213-218.	
VECTOR: PBAC108L	
FEATURES	Location/Qualifiers
source	1..188833
	/organism="Homo sapiens"
	/mol_type="genomic DNA"
	/db_xref="taxon:9606"
	/chromosome="22"
	/map="q13.2-13.3"
	/clone="CTA-268H5"
	/clone_lib="CIT978SK-A1"
ORIGIN	
Query Match 8.6%; Score 259.2; DB 9; Length 188833;	
Best Local Similarity 47.4%; Pred. No. 7,1e-58;	
Matches 1296; Conservative 0; Mismatches 1378; Indels 60; Gaps 15;	
QY 296	TCATAACAGTGCCTATCAGCCGGGTGCAGTGGCTAGTGCCTGAATATCCACACTTTGGG 355
Db 73238	TCAGAAACATAGCCGAGGCTGGGCACAGTGGCTCACATTTATATATCCAGCACTTTGGG 73297
QY 356	AGCCGGAGTGGCTGGATCACTTGTAGTCCAGGTTTCGAGCCAGCTGGCCCAACATGCT 415
Db 73298	AGGCTGAGCATGTGGATCACCTGAGGTCAGAGTTCAAGACCAGCTGGCCACACATAGT 73357
QY 416	GAACCCCTGCTCTACTAAAAATATAAAATTTAGCTGGGCATGTGGTGTGCTACCTGTAA 475
Db 73358	GAACCCCGTCTCTACTAAATATACAAATTTAGCCAGTGTAGGGTGCACACCTGTAG 73417
QY 476	TCCAGATACCTTGGAGGTTGAGGCGAGGAATCGCTTGAACCCGGGAGGCGAGTTGCC 535
Db 73418	TCCAGCTACTCAGGAGGCTGAGGCGAGGAATCACTTGAACCTGGGAGGCGAGTTGC 73477
QY 536	AGTGAACCAAGACTGTCACCTGCCTCCAGTCCGCAACAGAGTGAAGCTCCATCTCA 595
Db 73478	AGTGAACCGAGTGGCTCCACTGCACCTCCAGCTGGGCAACAGAGTGAATTCGCGCTCA 73537
QY 596	AAACAAACAAACAAACAGCTGCCATCATGTAGGATTCAGTGAATTCAGTGAAGACTGAG 655
Db 73538	AAACAAACAAACAAATTTAGAAACGTGAGTGGGTGAGTGCAG----- 73582
QY 656	CCTTGTGGAAGTGAAGCACTCACTAATCACCAGTTGTAGTATCAGTGATTAACCAAT 715
Db 73583	---TTTCTGTGCGCCCGTAACCACTCACTACAACTGCTGGCTTTAAAGCCACACAT 73639
QY 716	GATCCAGGTAAAGCCCTGAGGGTTTCAGAAAGATGCCGAGCGCTTTCAAGGTGCTGGGA 775
Db 73640	TTATTCCTCACAAGTTCTAGAACCAAGAGTCAAGAACCAAGGCATCCATGGGGCTGTC 73699
QY 776	TTGGTGGGCAAGCCCTCGAATATAGAAACAGTTCTCTGTATPACAAACAGAAAGCAGG 835
Db 73700	TGCTCCAGAGGCTCCAGGGGAGGCGCTCCCTGCTTGCAGCTTCTGGTGGCTCCAG 73759
QY 836	GCCCATGCTGGTGTGCTGCCAGGAACCTCAGTAGTAACT-----AAGACAGCACCGGTGCTG 890
Db 73760	GCACTCCTTGGCTTGGCGGTGCACCACTCCAAATTTCTGCTCCTCATCTTTCAATGECCT 73819
QY 891	CTTCCCGCAGGCACCTAGGCCAGTGGGGAACAGACTCACCAACAGTCCCGCCAGAG 950
Db 73820	CCTCCTCTGTGCTCACTCGCACTCAGTCTCCCTGCTGCTCTTAAAGTTTCAGGCTTGGCC 73879
QY 951	TGTCAGGGCCAGATGGGGAAGCAGCGGAGAAAGGTTCAGGTGGGATGGGAGGGGTC 1010
Db 73880	AGGTGGGTGGCTCACGCTGTAACTCTAACACTTTGGGAGACCAAGCGGGTGGATCAC 73939
QY 1011	AGGGCAAGAGGGTCAGGGCCAGGCTGAGGAAGCCCTGGACTGTAGGAATTTAGAGGA 1070
Db 73940	TTGAGGCGCAGAGTTTGAACCAAGCTTGGCCATGATGGTGAACCCCGTCTCTACTAAA 73999
QY 1071	GGTACCTGACCCGGCAGTGTGGTGAAGAGATTCAGGAAGTCTTCTGGAAGAGAGGCT 1130

Db 74000	AGACAAAATAATTAGCTGGGTATGTGGTGCACGCCTATAATCCAGCTACTCAGGAGACT 74059
QY 1131	GTCCGAGCTGAGACTCATAAGATGAGTGGGAGGGTGTTCAGGCAGAAAGACAGCACC 1190
Db 74060	GAGCAGAAAGAAATCGCTTGTATCTGGGAGCGGAGGTGTGAGTGAGCAAGATCGTGCC 74119
QY 1191	TACAAAAGCATGACTTTTGAGAGAAACATTCATCAATCAACTGATGAATTTTCAGACTGG 1250
Db 74120	ACCGACTCCAGCCTGGGTGACAGATGAGACTCCTCTCTCAAAAAAATAAATAATTC 74179
QY 1251	GCAGCT-----GGCTCATGCCCTGTATATCCAGCATCTTTGGAAAGGCTGAATGGG 1300
Db 74180	AGGCTCACACCTGTATAGGCTCACAGCTATATTTCCAGCACTTTGGGAGGCCACGGCAGG 74239
QY 1301	AGGATGACTTGAGCCTAGGCATTTGTGACAAAGCCCTGGGCAACATGTCAGACACCTGCCTC 1360
Db 74240	CAGATCACTGAGGTCAGGGGTTGAGACCCGGCTTGGCCCAACATGTTGAATGCTGTCTC 74299
QY 1361	CACAAAACAAACAAACAAACAAATCATTTATACCTGCTGATCCATCGGTACCAG--GTAC 1418
Db 74300	TACTAAAAATACAA--AATTAGCTGGGCATGTGGCAGTGCCTGTAATCCAGCTGTTC 74358
QY 1419	ATAGAAATGACTCAGGCAGATATGCTCTCTCTACTGTGGGAGAGCGGCGCTTATAC 1478
Db 74359	GGGAGGCTGAGCGAGGAGATCGTTTGAACCCAGGAGCGGGGTTTACAGTGAGCCGAGA 74418
QY 1479	TGCAGTAAGACATAGAGGAGGGAATATATCTTAAATGAGAGTACAGATTTTGAGAG 1538
Db 74419	TTGCCCCACCCGACTCCAGCTTGGGCGAACAAGAAAGTGTCCCTCCCAAAAAAATA 74478
QY 1539	CAACACACAGGCGACAGGCATATGTACGAGGGTAAAGAGGGAATCAGGGAAGCTTCTCAG 1598
Db 74479	AAAGTTCACTTCACATTTGAATGAACCCGTAATCCAGGATGCTGTCA--CCTCAAGA 74536
QY 1599	AGAAGTGACATTTTAAGCCGGACATGAAGAGTAAGAGTTAGTTTACCAGGATGCGA 1658
Db 74537	GACTGAGCTTAATCACTCTGCAAGATTTTCTTTTCCAAATAAAGGTCACTCGCAATCACA 74596
QY 1659	TGGAAAGGGTGAAGATGATGAGCGCAGAGGGAATGCAAGGATCATAGCCCTAGACAGG 1718
Db 74597	GGTAGACATCTATTTTGAGGAAGTAGGGGACA----AATTCAACTCAGTCTGCAGC 74652
QY 1719	GATCCTGAGCCCTTTGAGGAAGTGAAGAGAACAGCCGAGTCGTAGTGGTTAAGTAAC 1778
Db 74653	CTCTCCACCACTTTTGTGTAGTACAGCATCTGTCTCTCTCTCATGATGGGTTTT 74712
QY 1779	AAAGCTGAAGCAGGGAATCCCTGCTCATGCGAGGCTGTGAGTCACTCAGAGTGT 1838
Db 74713	TCTCTGGGACCCACCTTGTGCCCATTTTCAGTTTACAGAGCTGAGCCAGGCTTGCCCA 74772
QY 1839	TTGGGCTTTTGTTCCTGGGAGCAGTCGATTTTAAAGCAGGGAACAGCTGTATTCAGAG 1898
Db 74773	TTAAATCAGAGCATCTCCCTCGCCATAGTAGTCCAGGATGGTCCAGTCAGGSCAA 74832
QY 1899	TTGGGAAGATCCTGTGGTTGCTGCCTTGAAGGGGATGAACCTGAGAGCTGAGGCCCGAGG 1958
Db 74833	TCCAGGACTCTGGATGTCTTTGACGTAAGGCCACTTAAGAGGACAGTGCAGCC--GCC 74890
QY 1959	TGATAGGAGGATCCAGGGTGTGAGGAGGCTGGAGGTCGCGGTGATGGACACAGGCT 2018
Db 74891	ATCTTTCCAGCAGCCATCTTTGGTAGGCGACCTGGCTGAGATGGAGTCCACACAG-- 74948
QY 2019	GGGGCCAGGGATGGGAGGAGAGTAAATTTGGGAGAGGCTTGGGCTCTGGCCGAGGAA 2078
Db 74949	GAGAACGGCCAGAGATGGAAGTGAAGCCAGCTTTCAGTTCCGGAGCCAATACAC 75008
QY 2079	TGGATGGTGGCTGAACACAGGAGGAGAGATGCTTTAGGCCACTTTTGGNACACAGTAGG 2138
Db 75009	CCCTTTTCTGCTGAAGCCAGCTGAGAGGGGTTTCTGGCCCTTACTCATTTGTAGGTTTT 75068
QY 2139	GGAAGCAGGAGACACCCNAGGGAGTGCCTCAAGAGACACAGCAGGCTGGCATTTGA 2198
Db 75069	GAAGGAAGCAAGGTTTAAAGGAAAAGACACACACACACACACACACACACACAC 75128





```
repeat_region 1820..1935 /rpt_family="MIR3"
repeat_region 2081..2168 /rpt_family="L2"
repeat_region 3186..3264 /rpt_family="MIR"
repeat_region complement(3330..3398) /rpt_family="L2"
repeat_region 4043..4121 /rpt_family="MIR"
repeat_region complement(4624..4935) /rpt_family="AluY"
repeat_region complement(5045..5357) /rpt_family="AluY"
repeat_region 5686..5976 /rpt_family="AluY"
repeat_region complement(6115..6239) /rpt_family="MIR"
repeat_region 6448..6563 /rpt_family="MIR"
repeat_region complement(6914..6935) /rpt_family="(CG)n"
repeat_region 6935..6969 /rpt_family="(CA)n"
repeat_region complement(7330..7489) /rpt_family="FRAM"
repeat_region complement(7506..7679) /rpt_family="AluSq"
repeat_region 7698..7867 /rpt_family="L2"
repeat_region 8785..8824 /rpt_family="L2"
repeat_region 9672..9710 /rpt_family="(CAG)n"
repeat_region 9827..9909 /rpt_family="MIR3"
repeat_region 10004..10142 /rpt_family="MIR3"
repeat_region 10208..10418 /rpt_family="L2"
repeat_region complement(11095..11242) /rpt_family="MIR"
repeat_region 11593..11747 /rpt_family="L2"
repeat_region complement(12449..12492) /rpt_family="MIR"
repeat_region 12643..13021 /rpt_family="MSTB"
repeat_region complement(13389..13698) /rpt_family="AluSx"
repeat_region 13916..14570 /rpt_family="L2"
repeat_region complement(14794..14983) /rpt_family="MIR"
repeat_region 15256..15352 /rpt_family="L2"
repeat_region complement(15621..15765) /rpt_family="MIR"
repeat_region 15882..16007 /rpt_family="FLAN_A"
repeat_region 16015..16313 /rpt_family="AluY"
repeat_region complement(16317..16621) /rpt_family="AluSc"
repeat_region 16922..17267 /rpt_family="MLT1A0"
repeat_region complement(17338..17380) /rpt_family="MIR3"
repeat_region 17716..17968 /rpt_family="LTRL6C"
repeat_region complement(18551..18665) /rpt_family="MIR3"
repeat_region 18724..18877

/rpt_family="MIR"
18913..19052
/rpt_family="L2"
complement(19781..19868)
/rpt_family="MIR"
20196..20358
/rpt_family="L2"
complement(20898..20999)
/rpt_family="MIR3"
22083..22103
/rpt_family="GC rich"
complement(22173..22290)
/rpt_family="MIR"
22787..23087
/rpt_family="AluSx"
complement(23387..23632)
/rpt_family="AluSc"
complement(23851..24293)
/rpt_family="L2"
complement(24321..24432)
/rpt_family="L2"
25322..25376
/rpt_family="MIR"
complement(25519..25592)
/rpt_family="MIR3"
25899..25978
/rpt_family="CT-rich"
complement(26046..26198)
/rpt_family="MIR"
26964..27092
/rpt_family="L2"
27149..27383
/rpt_family="AluSx"
complement(27391..27552)
/rpt_family="AluJb"
27553..27628
/rpt_family="AluSx"
27630..27872
/rpt_family="L1MC4a"
complement(27889..28085)
/rpt_family="L1MC4a"

Query Match 8.6%; Score 258; DB 9; Length 39634;
Best Local Similarity 65.1%; Pred. No. 1.5e-57;
Matches 396; Conservative 0; Mismatches 210; Indels 2; Gaps 1;

QY 2394 GCCAGGTCGGTGCCTCAGCTTCCAGCTTCCATCTAGCAGCTTGGGAGGCTGAGGTGGGTGGAT 2453
DB 35519 GCCAGGCTTGGTGGCTCAGCCTGTAAATCCAGCACTTTGGGAGGCAAGCGGGCGGAT 35578

QY 2454 TACCTCAGGTCAGGAGTTCAAGACCAGCCTAGACAGTAGAGAGAAACCCCATCTCTACTG 2513
DB 35579 CACCTGAGTTCGGAGTTGGAGACCAGCCTCACCAACATGGAGAGAACCCCATCTCTACTA 35638

QY 2514 AAGATATAAAATTAGTGGGCTAGTGGCATATGCTGTAAATCCAGCTAGTGGGAAGC 2573
DB 35639 AAAATACAAAATTAGCCAGGCATGTTGGCACATGCTGTAAATCCAGCTACTCGGAGGC 35698

QY 2574 TGAGGCGAGGAATCCGTTGAACCCGAGGAGGAGGTTGCGGTGAGCTGAGATCGTGCC 2633
DB 35699 TGAGGCGAGGAATTCGTTGAACCCGAGGAGGTTGCGGTGAGCTGAGATCATGCC 35758

QY 2634 ATTGCACCTCAGCCTGGGTGACAAAGCAAGACCTCGTCTCAATAATAATAATAATACAA 2693
DB 35759 ATTACACTCCAGCCTCAGCAAGAGAGTGAACCTCGTCTCAAAAAATAAATAATATAT 35818

QY 2694 AACAGAGAGCTGGTGCATCCAGCTACCTACTTTTCAGGAGATGTACTCCCTTACC 2753
DB 35819 TTGTGTTAGGCATGGTGACTCAGGCTTACCTGTAAATCCCTAGAACTT--TGGGAGGCC 35876

QY 2754 CRAGGGCAAGGATGGAGAACCAAGTTTGATTATGATTTTATGAGCACCCTACTAGTCC 2813
DB 35877 GAGGTAGGAGGATGCTTTGAGGCCAGGAGTTCAAGACTAGCCTGGGCAACATAGCAAGAT 35936
```

```

Qy 2814 TCATCCCTGGCTAGCTGGAATGACTCAGATGAGCTGAAGAGTCCCTCAGGAA 2873
Db 35937 CCCATCTCTACAGAAAAGTAATTTTAAATAATATGACATTTTGAGCTGATGATGAAA 35996
Qy 2874 CCTCAGTAAAGAGGAGGAATCGCGCGCGGCTGCTCAGCCTGTGTAATCCCAACA 2933
Db 35997 AAAATTTTAAAGAAAGAAAGAGGCGCGCGGCTGCTCAGCCTGTGTAATCCCAACA 36056
Qy 2934 CTTTGGGAGGCTGAGTGGTGGATCACAAGTCAAGGATCGAGACCATCTGGCTTAAC 2993
Db 36057 CTTTGGGAGGCTGAGTGGTGGATCATGAGGTCAAGGATCAAGACCATCTGGCTTAAC 36116
Qy 2994 ACAAGTAA 3001
Db 36117 GCGGTGAA 36124

RESULT 9
AC140106/c 40948 bp DNA linear PRI 21-FEB-2003
LOCUS Homo sapiens chromosome 17, clone XXfos-801239H9, complete
DEFINITION
ACCESSION AC140106
VERSION AC140106.1 GI:28460977
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 40948)
AUTHORS Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N.,
Anderson, S., Arachchi, H. M., Barna, N., Bastien, V., Bloom, T.,
Boguslavskiy, L., Boukhalter, B., Camarata, J., Chang, J., Choepel, Y.,
Collymore, A., Cook, A., Cooke, P., Corum, B., DeArellano, K., Diaz, J. S.,
Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S.,
Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S.,
Graham, L., Grand-Fierre, N., Hafez, N., Hagopian, D., Hagos, B.,
Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R.,
Lindblad-Toh, K., Liu, G., Lui, A., Mabbitt, R., MacLean, C.,
Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M.,
Meldrim, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J.,
Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P.,
O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N.,
Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P.,
Roman, J., Schauer, S., Schuback, R., Seaman, S., Severy, P., Smith, C.,
Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J.,
Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H.,
Venkataraman, V. S., Viel, R., Vo, A., Wilson, B., Wu, X.,
Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (21-FEB-2003) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
All repeats were identified using RepeatMasker:
Smit, A. F. A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Web site: http://www-seq.wi.mit.edu
Center code: WIBR
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L29233
Center clone name: 801239_H_9
-----
FEATURES
Location/Qualifiers
source 1..40948
/organism="Homo sapiens"

```

```

/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="17"
/map="17"
/clone="XXfos-801239H9"
/clone_lib="Human Female Fosmid Library"
99..339
/rpt_family="MIR"
complement(379..509)
/rpt_family="MIR3"
819..929
/rpt_family="MIR3"
complement(1313..1573)
/rpt_family="L2"
complement(1789..1933)
/rpt_family="MIR"
complement(2078..2255)
/rpt_family="AluJo"
2256..2281
/rpt_family="(CAAA)n"
complement(2282..2402)
/rpt_family="AluJo"
complement(2424..2860)
/rpt_family="L2"
3067..3142
/rpt_family="GA-rich"
complement(3585..3716)
/rpt_family="MIR"
3854..3883
/rpt_family="(TC)n"
complement(3914..4252)
/rpt_family="Aluub"
complement(4559..4690)
/rpt_family="AluSg/x"
4705..4792
/rpt_family="LTR25"
complement(4793..4952)
/rpt_family="AluSg"
complement(4956..5265)
/rpt_family="AluSg"
5441..5605
/rpt_family="FRAM"
complement(5677..5982)
/rpt_family="AluY"
complement(5985..6276)
/rpt_family="AluSx"
6669..6805
/rpt_family="AluSg/x"
6807..7101
/rpt_family="AluSx"
7117..7365
/rpt_family="AluSx"
7368..7667
/rpt_family="AluSx"
complement(7938..8172)
/rpt_family="AluSg/x"
8409..8544
/rpt_family="MIR"
8594..8860
/rpt_family="AluSx"
complement(8994..9113)
/rpt_family="FLAM_C"
complement(9158..9464)
/rpt_family="AluSx"
complement(9472..9767)
/rpt_family="AluSg"
complement(9889..10199)
/rpt_family="AluSg1"
10422..10706
/rpt_family="AluSg"
10773..11083
/rpt_family="AluY"
complement(11296..11685)

```

Mon Sep 20 09:09:29 2004

```
repeat_region      /rpt_family="AluSg"
11823..11923
/rpt_family="MLT1K"
complement(11936..12253)
/rpt_family="AluJb"
complement(12268..12576)
/rpt_family="AluSc"
complement(12610..12817)
/rpt_family="MLT1K"
complement(12820..12988)
/rpt_family="AluSp"
complement(12990..13283)
/rpt_family="AluSg"
complement(13287..13418)
/rpt_family="MLT1K"
13488..13755
/rpt_family="AluJb"
13761..13790
/rpt_family="CA)n"
13794..13922
/rpt_family="AluSg/x"
13924..14060
/rpt_family="AluSg"
14061..14370
/rpt_family="AluSx"
14371..14534
/rpt_family="AluSg"
14541..14717
/rpt_family="MIR"
complement(15059..15160)
/rpt_family="MIR"
15649..15720
/rpt_family="MIR"
16036..16085
/rpt_family="MIR"
16109..16410
/rpt_family="AluSg"
complement(16447..16654)
/rpt_family="MER3"
17064..17245
/rpt_family="L1"
complement(17291..17399)
/rpt_family="MIR3"
complement(17799..17909)
/rpt_family="LIME3A"
18318..18373
/rpt_family="GC rich"
complement(18508..18633)
/rpt_family="FLAM_A"
complement(18636..18797)
/rpt_family="LIME3A"
complement(18798..18869)
/rpt_family="MLT1B"
18872..19003
/rpt_family="AluJo/FLAM"
19005..19294
/rpt_family="AluSx"
complement(19586..19802)
/rpt_family="U3"
20084..20410
/rpt_family="AluJo"
complement(20415..20784)
/rpt_family="LIMEB4"
complement(20785..21046)
/rpt_family="AluJb"
complement(21048..21102)
/rpt_family="L1M4"
21103..21398
/rpt_family="AluSg"
complement(21399..21594)
/rpt_family="L1M4"
21595..21903
/rpt_family="AluSg"
```

```
repeat_region      complement(21904..22221)
Query Match        8.6%; Score 258; DB 9; Length 40948;
Best local Similarity 65.1%; Pred. No. 1.5e-57;
Matches 396; Conservative 0; Mismatches 210; Indels 2; Gaps 1;
QY 2394 GGCAGGTGGGTGCTCAGCGTTGCCATTCTAGCACCTTTGGGAGGCTGAGGTGGGTGAT 2453
Dd 25335 GCCAGGCTTGGTGGCTCAGCGCTGTATCCAGCACCTTTGGGAGGCCAAGCGGCGGAT 25576
QY 2454 TACCTCAGGTGAGGATTCAAGACCCAGCTAGACAACTAGAGAGAAACCCATCTCTACTG 2513
Dd 25575 CACCTGAGGTGGGAGTTGGAGACAGCCTCACCACATGGAGAAACCCATCTCTACTA 25516
QY 2514 AAGATATAAATTTAGCTGGCGTAGTGCATATGCTGTATCCAGCTAGTCCGGAGGC 2573
Dd 25515 AAAATACAAAATTAGCCAGGCATGGTGCAATGCCCTGTATCCAGCTATCTGGAGGC 25456
QY 2574 TGAGGCAGGAGATCGCTTTGAACCCGAGAGCAGAGTTGCGGTGAGCTGAGATCGTGCC 2633
Dd 25455 TGAGGCAGGAGATTGCTTTGAACCCGAGAGTTGCGGTGAGCTGAGATCATGCC 25396
QY 2634 ATTCACTCCAGCTGGGTGACAAAGCAAGACCTGCTCTCAATATAATAATAATACAA 2693
Dd 25395 ATTCACTCCAGCTGGAGCAAGAGTGAACCTCCGTCTCAAAAATAAAAATATATAT 25336
QY 2694 AACAGAGGAGCGCTGGGTCAATCCAGCTACCTACTTTTTCAGGAGAAATGCTCTCCCTTACC 2753
Dd 25335 TTGTGTTAGGATGGTGAATCAGCTTACAGCTGTATCTAGAACTT--TGGGAGGCC 25278
QY 2754 CAAGGCAAGGATGGGAGAACCCAGTTTGTATATGTCATTTATTTAGCACCTTCTAGTACC 2813
Dd 25277 GAGGTAGGAGATTGCTTTGAGGCCAGGAGTTTCAAGACTAGCTGGGCAACATAGCAAGAT 25218
QY 2814 TCATCCCTGGGTAGGCTGGATGAGCTCAGATGGAGCTGAGAGTCCCTCCCTCAGGAA 2873
Dd 25217 CCCATCTCTACAGAAAGTAAATTTTAAATAATATGACATTTTGGCTGCAATGATGAA 25158
QY 2874 CTTCACTAGAAGAGGAGGAAATCGCGGGCGGTGGCTCACCGCTGTATATCCCAACA 2933
Dd 25157 AAAATTTAAAGAAAGAAAGAGGCGCGGCGGCGGCTGATATCCAGCA 25098
QY 2934 CTTTGGGAGGCTGAGGTGGTGGATCAAGGTGAGGATCGAGACATCTCTGGCTAAC 2993
Dd 25097 CTTTGGGAGGCTGAGGTGGGCGGATCATGAGGTGAGGATCAAGACCATCTCTGGCTAAC 25038
QY 2994 ACAGTGAA 3001
Dd 25037 GCGGTGAA 25030
```

```
RESULT 10
AC144507          41867 bp      DNA      linear      HTG 25-APR-2003
LOCUS             Homo sapiens chromosome 17 clone XXfos-80066E9 map 17, ***
DEFINITION        SEQUENCING IN PROGRESS ***, 2 ordered pieces.
ACCESSION         AC144507
VERSION           AC144507.1 GI:30103061
KEYWORDS          HTG; HTGS_PHASE2; HTGS_FULLTOP; HTGS_ACTIVEFIN.
SOURCE            Homo sapiens
ORGANISM          Homo sapiens
REFERENCE          1 (bases 1 to 41867)
AUTHORS           Birren,B., Nusbaum,C. and Lander,E.
TITLE             Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
JOURNAL           Eukaryota; Euthera; Primates; Catarrhini; Hominidae; Homo.
REFERENCE         2 (bases 1 to 41867)
AUTHORS           Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
                  Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,
                  Boguslavskiy,L., Boukhalter,B., Camarata,J., Choepel,Y.,
                  Collymore,A., Cook,A., Cooke,P., Corum,B., DeArellano,K.,
                  Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,
```

Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kellis, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Lui, A., Mabbitt, R., Maclean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Medlir, J., Meus, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Nibhu, C., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schupbach, R., Seaman, S., Seaver, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, C., Testay, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkatachalam, V.S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission

Submitted (25-APR-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

All repeats were identified using RepeatMasker:

Smith, A.P.A. & Green, P. (1996-1997)

<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: W1BR

Web site: <http://www-seq.wi.mit.edu>

Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)

Project Information

Center project name: L29228

Center clone name: 80066\_E\_9

NOTE: This is a 'working draft' sequence. It currently consists of 2 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.

This sequence will be replaced

by the finished sequence as soon as it is available and the accession number will be preserved.

1 35848: contig of 35848 bp in length

35949 41867: gap of 100 bp

35949 41867: contig of 5919 bp in length.

# FEATURES

source

Location/Qualifiers

1. .41867

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="taxon:9606"

/chromosome="17"

/map="17"

/clone="XXfos-80066E9"

/clone\_lib="Human Female Fosmid Library"

ORIGIN

Query Match 8.6%; Score 258; DB 2; Length 41867;

Best Local Similarity 65.1%; Pred. No. 1.5e-57;

Matches 396; Conservative 0; Mismatches 210; Indels 2; Gaps 1;

QY 2394 GGCGAGGTCGGTGCACGCTTGCCATCTAGCACTTTGGGAGGCTGAGGTGGTGGAT 2453

DB 5394 GCAGGCTTGGTGGCTCAGCTTGTAATCCAGCACTTTGGGAGGCAAGCGGGCGGAT 5453

QY 2454 TACCTCAGTTCAGGAGTTCAAGACCAAGCCTAGACACGTAGAGAACCCCATCTCTACTG 2513

DB 5454 CACCTGAGTTCGGAGTTGGAGCAGCCTCCACACATGGAGAACCCCATCTCTACTA 5513

QY 2514 AGATATAAAATAGTGGGGGTAGTGGCATATGCTGTAATCCAGCTAGTCGGGAAGC 2573

DB 5514 AAAATACAAAATAGCCAGGCATGTGGCACATGCTGTAATCCAGCTACTCGGAGGC 5573

QY 2574 TGAGGCGAGGAATCGCTTGACCCGAGAGGAGGAGGTTGGCGTGGAGCTGAGATCGTGCC 2633

DB 5574 TGAGGCGAGGAATTCCTTGAACCCGGGAGGTGGAGGTTGGCGTGGAGCTGAGATCATGCC 5633

QY 2634 ATTGCACCTCCAGCCTCGGTGACAAAGCAAGACCTCGTCTCAATAATAATAATTACAA 2693

DB 5634 ATTACACTCCAGCCTGAGCAAGAGAGTGAAACTCCGCTCTCAAAAAATAAAATATAT 5693

QY 2694 AACAGAAGGAGCCTCGGTGATCCCGAGCTACCTCTTTTCAGGAGATGTAATCCCTTACC 2753

DB 5694 TTGTGGTTAGGCATGGTGCATCAGGCTTACACCTGTGAATCCTAGAACTT--TGGAGAGCC 5751

QY 2754 CAAGGGCAAGGATGGGAGAACCAAGTTTGATTTATGCAATTTTATGAGCACCTACTGAGTCC 2813

DB 5752 GAGGTAGGAGGATGCTTGGAGCCAGAGTTCAAGACTAGCTGGGCAACATAGCAAGAT 5811

QY 2814 TCATCCCTGGGTAGGCTGGAATGGAATCAGATGGAGCTTCAAGAGTCCCTCAGGAA 2873

DB 5812 CCCATCTCTACAGAAAAAGTAAATTTTAAATAATTTATGACATTTTTCAGCTGATGATAA 5871

QY 2874 CCTCACTAGAAAAAGAGGAGGAATCGGCGCGGGTGGCTCAGCGCTGTAATCCACACA 2933

DB 5872 AATATTTAAAGAAAGAAAGAGAGGGCGGGCGGTGGCTCAGCGCTGTAATCCAGCA 5931

QY 2934 CTTTGGGAGCTGAGGTGGTGGATCAAGAGTCAAGGATCAGAGATCAGAGACCATCTCGCTAAC 2993

DB 5932 CTTTGGGAGCTGAGGTGGCGGATCATGAGTCAGGAGATCAAGACCATCTCGCTAAC 5991

QY 2994 ACAGTGAA 3001

DB 5992 GCGGTGAA 5999

RESULT 11

HSJ9617/c

LOCUS

DEFINITION

40M01, complete sequence.

ACCESSION

AJ009617 AL035361

VERSION

AJ009617.5 GI:29466470

KEYWORDS

HTG.

SOURCE

Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1

AUTHORS

Sudbrak, R., Mueller, I., Kosiura, A., Radelof, U., Hennig, S., Ramser, J., Francis, F., Steffens, C., Klein, M., Seranski, P., Poustka, A., Lehrach, H. and Reinhardt, R.

JOURNAL

Unpublished

REFERENCE

2 (bases 1 to 153609)

AUTHORS

MPIMG.

TITLE

Direct Submission

JOURNAL

Submitted (17-JUL-1998) MPIMG, Abt. Lehrach, Max Planck Institut fuer Molekulare Genetik, Ihnestrasse 73, Berlin, 14195, Germany

COMMENT

On Apr 1, 2003 this sequence version replaced gi:9211531.

contig 01 1. .153609

Clone received from the Resource Centre of the Human Genome Project at the Max-Planck-Institute for Molecular Genetics.

FEATURES

Location/Qualifiers

1. .153609

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="taxon:9606"

/chromosome="17"

/map="17p11.2"

/clone="COSMID 40M01"

/clone\_lib="RPC11.3-5 Human PAC library, originating institute: Roswell Park Cancer Institute, creator: Pieter de Jonge, P. Ioannou 40M01: L4/FS17 (human chromosome 17 specific cosmid library)"

/note="region between markers D17S842-D17S953"

1

misc\_feature

/note="Sp6\_end:RPI-149D14"

153609

misc\_feature

/note="T7\_end:COSMID 40M01"

ORIGIN







2934 CTTTGGAGGCTGAGTGGGTGATCATCAAGGTCAGAGATCGAGACCATCTGGCTAAC 2993  
|||||  
36852 CTTTGGAGGCTGAGTGGGTGATCATGAGTTCAGAGATCAAGACCATCTGGCTAAC 36911  
|||||  
2994 ACAGTGA 3001  
|  
36912 GCGGTGA 36919  
|

RESULT 13  
AC007952/c  
LOCUS Homo sapiens chromosome 17, clone RP11-160E2, complete sequence.  
DEFINITION AC007952  
ACCESSION AC007952.15 GI:29501895  
VERSION HTG.  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Homo sapiens

REFERENCE  
AUTHORS Birren, B., Nussbaum, C. and Lander, E.  
TITLE Homo sapiens chromosome 17, clone RP11-160E2  
JOURNAL Unpublished  
AUTHORS 2 (bases 1 to 158103)  
REFERENCE Baker, J., Baldwin, J., Barna, N., Beckerly, R., Benn, J., Brown, A.,  
Caestele, A., Cerny, J., Colangelo, M., Collins, S., Collymore, A.,  
Cook, A., Cooke, P., DeArelano, K., Depayre, E., Devon, K., Dewar, K.,  
Donelan, L., Doyle, M., Ferreira, P., FitzHugh, W., Forrest, C.,  
Funke, R., Gage, D., Galagan, J., Gardyna, S., Gilbert, D., Grant, G.,  
Hagos, B., Heaford, A., Horton, L., Howland, J. C., Jones, C., Kann, L.,  
Karatas, A., Lechoczy, J., Lieu, C., Locke, K., Macdonald, P.,  
Marquis, N., McEwan, P., McGurk, A., McKernan, K., McLaughlin, J.,  
Meldrim, J., Molla, M., Morris, W., Morrow, J., Mychaleckyj, J.,  
Naylor, J., Niloff, M., O'Connor, T., O'Donnell, P., Pavlin, B.,  
Peterson, K., Pollara, V., Riley, R., Roberts, D., Roy, A., Severy, P.,  
Stange-Thomann, N., Stojanovic, N., Stone, C., Subramanian, A.,  
Tesfaye, S., Torruella-Miller, I., Vassiliev, H., Vo, A., Wagner, A.,  
Wheeler, J., Wu, X., Wyman, D., Ye, W. J. and Zody, M.

Direct Submission  
Submitted (02-JUL-1999) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
3 (bases 1 to 158103)  
REFERENCE  
AUTHORS Birren, B., Nussbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S.,  
Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhalter, B.,  
Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A.,  
Cook, A., Cooke, P., DeArelano, K., Dewar, K., Diaz, J. S., Dodge, S.,  
Faro, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J.,  
Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hafez, N.,  
Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,  
Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R.,  
Lindblad-Toh, K., Liu, G., MacLean, C., Macdonald, P., Major, J.,  
Mathews, C., McCarthy, M., Meldrim, J., Meneus, L., Mihova, T.,  
Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C.,  
Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J.,  
Peterson, K., Phunkhang, P., Pierre, N., Raymond, C., Retta, R.,  
Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schuback, R.,  
Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N.,  
Stojanovic, N., Talamas, J., Tesfaye, S., Theodore, J., Topham, K.,  
Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X.,  
Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission  
Submitted (20-NOV-2002) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
4 (bases 1 to 158103)  
REFERENCE  
AUTHORS Birren, B., Nussbaum, C., Lander, E., Abouelleil, A., Allen, N.,  
Anderson, S., Arachchi, H. M., Barna, N., Bastien, V., Bloom, T.,  
Boguslavskiy, L., Boukhalter, B., Camarata, J., Chang, J., Choepel, Y.,  
Collymore, A., Cook, A., Cooke, P., Corum, B., DeArelano, K.,  
Diaz, J. S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S.,  
Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S.,

Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B.,  
Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,  
Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R.,  
Lindblad-Toh, K., Liu, G., Lui, A., Mabbitt, R., MacLean, C.,  
Macdonald, P., Major, J., Manning, J., Menga, V., Murphy, T., Naylor, J.,  
Meldrim, J., Meneus, L., Mihova, T., Mlenga, V., O'Donnell, P.,  
Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Neill, P.,  
O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N.,  
Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P.,  
Roman, J., Schauer, S., Schuback, R., Seaman, S., Severy, P., Smith, C.,  
Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M.,  
Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M.,  
Vassiliev, H., Venkataraman, V. S., Viel, R., Vo, A., Wilson, B., Wu, X.,  
Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission  
Submitted (03-APR-2003) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Apr 3, 2003 this sequence version replaced gi:25140154.  
All repeats were identified using RepeatMasker:  
Smit, A. F. A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html  
-----  
Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: http://www-seq.wi.mit.edu  
Contact: sequence.submissions@genome.wi.mit.edu  
-----  
Project Information  
Center project name: L913  
Center clone name: 160\_E\_2  
-----

This clone contains two copies of a large inverted duplication from  
about base pair 55,000 to 82,000 and about base pair 130,000 to  
158,000. Bases 154,031 to 154,092 are represented here by the  
corresponding bases of the first copy of the duplication.

FEATURES  
source  
1. 158103  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="17"  
/map="17"  
/clone="RP11-160E2"  
/clone\_lib="RP11-11 Human Male BAC"  
repeat\_region 331..374  
/rpt\_family="CCCCCA)n"  
repeat\_region 2035..2157  
/rpt\_family="MIR3"  
repeat\_region 2607..2722  
/rpt\_family="MIR3"  
repeat\_region 2868..2955  
/rpt\_family="MIR"  
repeat\_region 3973..4051  
/rpt\_family="MIR"  
repeat\_region complement(4117..4185)  
/rpt\_family="L2"  
repeat\_region 4830..4908  
/rpt\_family="MIR"  
repeat\_region complement(5411..5722)  
/rpt\_family="AluY"  
repeat\_region complement(5832..6144)  
/rpt\_family="AluY"  
repeat\_region 6473..6763  
/rpt\_family="AluY"  
repeat\_region complement(6902..7026)  
/rpt\_family="MIR"  
repeat\_region 7235..7350  
/rpt\_family="MIR"  
repeat\_region 7695..7762  
/rpt\_family="(CA)n"  
repeat\_region complement(8123..8282)  
/rpt\_family="FRAM"  
repeat\_region complement(8299..8472)



Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,  
O'Neill, D., Olivari, T.M., Oliver, J., Peterson, K., Pierre, N.,  
Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,  
Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,  
Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,  
Tessier, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J.,  
Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J.,  
Young, G., Zainoun, J., Zimmer, A. and Zody, M.

Direct Submission  
Submitted (21-APR-2000) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
3 (bases 1 to 176584)

REFERENCE  
AUTHORS  
Birren, B., Linton, L., Nusbaum, C., Lander, E., Abram, H., Allen, N.,  
Anderson, S., Baldwin, J., Barna, N., Bastien, V., Beda, F.,  
Boguslavskiy, L., Boukhgalter, B., Brown, A., Burkett, G.,  
Campopiano, A., Castle, A., Choepel, Y., Colangelo, M., Collins, S.,  
Collins, A., Cooke, P., DeArillano, K., Dewar, K., Diaz, J.S.,  
Dodge, S., Domino, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D.,  
Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,  
Grand-Pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L.,  
Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karakas, A.,  
Klein, J., LaRocque, K., Lamazares, R., Landers, T., Lehoczy, J.,  
Levine, R., Lieu, C., Liu, G., Locke, K., Macdonald, P., Marguis, N.,  
McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McPheeters, R.,  
Meidrim, J., Meneus, L., Mihova, T., Miranda, C., Mlenga, V., Morrow, J.,  
Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,  
O'Neill, D., Olivari, T.M., Oliver, J., Peterson, K., Pierre, N.,  
Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,  
Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,  
Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,  
Tessier, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J.,  
Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J.,  
Young, G., Zainoun, J., Zimmer, A. and Zody, M.

Direct Submission  
Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On May 22, 2000 this sequence version replaced gi:7630749.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: <http://www.seq.wi.mit.edu>  
Contact: [sequence.submissions@genome.wi.mit.edu](mailto:sequence.submissions@genome.wi.mit.edu)  
----- Project Information  
Center project name: L9984  
Center clone name: 297\_C\_16  
----- Summary Statistics  
Sequencing vector: M13; M77815; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.960731  
Consensus quality: 159055 bases at least Q40  
Consensus quality: 167576 bases at least Q30  
Consensus quality: 170993 bases at least Q20  
Insert size: 165000; agarose-ff  
Quality coverage: 4.1 in Q20 bases; agarose-ff  
Quality coverage: 3.9 in Q20 bases; sum-of-contigs

----- NOTE: This is a 'working draft' sequence. It currently  
\* consists of 38 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 1405: contig of 1405 bp in length  
\* 1406 1505: gap of 100 bp  
\* 1506 3162: contig of 1657 bp in length  
\* 3163 3262: gap of 100 bp  
\* 3263 4753: contig of 1491 bp in length

4754 4853: gap of 100 bp  
\* 4854 6191: contig of 1338 bp in length  
\* 6192 6291: gap of 100 bp  
\* 6292 7658: contig of 1367 bp in length  
\* 7659 7758: gap of 100 bp  
\* 7759 9247: contig of 1489 bp in length  
\* 9248 9347: gap of 100 bp  
\* 9348 10791: contig of 1444 bp in length  
\* 10792 10891: gap of 100 bp  
\* 10892 13201: contig of 2310 bp in length  
\* 13202 13301: gap of 100 bp  
\* 13302 15490: contig of 2189 bp in length  
\* 15491 17422: contig of 1832 bp in length  
\* 17423 19981: contig of 2459 bp in length  
\* 19982 20081: gap of 100 bp  
\* 20082 21487: contig of 1406 bp in length  
\* 21488 21587: gap of 100 bp  
\* 21588 23066: contig of 1479 bp in length  
\* 23067 23166: gap of 100 bp  
\* 23167 25655: contig of 2489 bp in length  
\* 25656 27996: gap of 100 bp  
\* 27997 28096: contig of 2241 bp in length  
\* 28097 30125: contig of 2029 bp in length  
\* 30126 30225: gap of 100 bp  
\* 30226 33060: contig of 2835 bp in length  
\* 33061 33160: gap of 100 bp  
\* 33161 37147: contig of 3987 bp in length  
\* 37148 40807: contig of 3560 bp in length  
\* 40808 40907: gap of 100 bp  
\* 40908 43945: contig of 3038 bp in length  
\* 43946 44045: gap of 100 bp  
\* 44046 46338: contig of 2293 bp in length  
\* 46339 50865: contig of 4427 bp in length  
\* 50866 50965: gap of 100 bp  
\* 50966 56185: contig of 5220 bp in length  
\* 56186 56285: gap of 100 bp  
\* 56286 61204: contig of 4919 bp in length  
\* 61205 61304: gap of 100 bp  
\* 61305 66924: contig of 5620 bp in length  
\* 66925 70244: gap of 100 bp  
\* 70245 70900: contig of 3876 bp in length  
\* 70901 76151: contig of 5151 bp in length  
\* 76152 76251: gap of 100 bp  
\* 76252 81819: contig of 5568 bp in length  
\* 81820 87341: contig of 5422 bp in length  
\* 87342 87441: gap of 100 bp  
\* 87442 92459: contig of 5017 bp in length  
\* 92459 92559: gap of 100 bp  
\* 92559 98672: contig of 6114 bp in length  
\* 98673 98772: gap of 100 bp  
\* 98773 103721: contig of 4949 bp in length  
\* 103722 103821: gap of 100 bp  
\* 103822 112632: contig of 8811 bp in length  
\* 112633 121203: contig of 8471 bp in length  
\* 121204 121303: gap of 100 bp  
\* 121304 127699: contig of 6396 bp in length  
\* 127700 127799: gap of 100 bp  
\* 127800 140599: contig of 12800 bp in length  
\* 140600 140699: gap of 100 bp  
\* 140700 156652: contig of 15953 bp in length  
\* 156653 156752: gap of 100 bp  
\* 156753 176584: contig of 19832 bp in length.

Location/Qualifiers  
1. .176584  
/organism="Homo sapiens"

FEATURES  
source

/mol_type="genomic DNA" /db_xref="taxon:9606" /chromosome="22" /map="22" /clone="RP11-297C16" /clone_lib="RP11-11 Human Male BAC" 1.1405 /note="assembly_fragment" 1506.3162 /note="assembly_fragment" 3263.4753 /note="assembly_fragment" 4834.6191 /note="assembly_fragment" 6292.7658 /note="assembly_fragment" 7159.9247 /note="assembly_fragment" 9348.10791 /note="assembly_fragment" 10892.13201 /note="assembly_fragment" 13302.15490 /note="assembly_fragment" 15591.17422 /note="assembly_fragment"		Query Match Best Local Similarity 47.4%; Pred. No. 1.5e-57; Length 176584; Matches 1295; Conservative 0; Mismatches 1380; Indels 59; Gaps 15;	
misc_feature	296	TGATAACAGTGCCTCAGCGGGTGCAGTGGCTAGTGCCTGAAATCCCAACTTTGGG	355
Db	128084	TCAGAAACATGAGCGAGGCTGGGCACAGTGGCTCACATTATATCCAGCACTTTGGG	128143
QY	356	AGCGGAGGTGGGTGATCACTTGAAGTTCAGGAGTTCGAGACGAGCTGCCAACATGGT	415
Db	128144	AGGCTGAGGCATGTGGATCACCTGAGTCAAGTTCAGGAGTTCAGACCGAGCTGCCAACATAGT	128203
QY	416	GAACCCCTCTCTACTATAAAATATAAAATATAGCTGGGATGCTGGGATGCTGTTACCTGTAA	475
Db	128204	GAACCCCTCTCTACTATAAAATATAAAATATAGCAGGAGTTCAGGAGTTCAGACCTGTAG	128263
QY	476	TCCAGATACCTTGGGAGTTCAGGAGGAGTTCGCTTGAACCCGAGGAGCAGATGTC	535
Db	128264	TCCAGTACTCAGGAGGCTGAGGAGGAGTTCATCTTGAACCTGGAGGAGGAGTTC	128323
QY	536	AGTGACCAAGACTGTCCTGCACTCCAGTTCGGGCAACAGAGTGAAGCTCCATCTCA	595
Db	128324	AGTGAACGAGGTGGCTCCACTGCACTCCAGCTGGGCAACAGAGTGAAGTCCGCTCA	128383
QY	596	AACAAACAAACAAAGCAGTCCCATCATGTAGGATTCAGTGAAGTGAAGTGAAGTGAAG	655
Db	128384	AAACAAACAAACAAATATAGAAACGTGAGCTGGGTGAGTGCAG-----	128428
QY	656	CTTTGTGAAAGTGAAGTCACTACTAATCAACAGGTGTAGTATCAGTATACCACTCAAT	715
Db	128429	---TTTCTGTGGCGCGGTACCACTCACTCAAACTGCTGGCTTAAGCCAAACAGAT	128485
QY	716	GATCAGGTAAAGCCCTGAGGGTTGAGAAAGTTCGCGGAGCCCTTCAAGGTGCTGGGA	775
Db	128486	TTATCTCTCAGAGTCTAGAGAGCAGAGTTCAGAAACCAAGGCATCCATGGGGTGTGC	128545
QY	776	TTGTGGGCAAGCCCTCGAATATAGAAACAGTTCCTGTATTAACAGAAAGCAGGAG	835
Db	128546	TGCCTCAGAGGCTCAGGGGAGGAGGCGTCTGCTCTTCAGAGTCTTGTGGGTCCAG	128605
QY	836	GCCCATGCTGGTGTGTCAGGAACTCAGTAGTAAT-----AAGACAGCACCGGTGTG	890
Db	128606	GCATCTCTGGCTTGGCTGACCACTCCATTTCTGCCTCCATCTCTTCAATGGCT	128665
QY	891	CTTCCCAAGCGCACTTAGCCAGTGGGAAACAGACTCACCACAGTCCCGCCAGAG	950
Db			
128666	CTCTCTGTGTCTACTGGGACTAGTCTCCCTGTCTCTTAAAGTTTCAAGGCTTGCC	128725	
QY	951	TGTTTCAAGGCAAGATGGGAGACCGGGGAGAAAGTTCAGGTTGGGAGGGGCTC	1010
Db	128726	AGTGGCGGTGGCTCAGCGCTGTAATCTTAAACACITTTGGGAGCAACGCGGTGATCAC	128785
QY	1011	AGGCAAGAGGGTTCAGGGCCAGGCTAGGGAGCCCTGGAGCTGTAGAAATTTAGAGA	1070
Db	128786	TTGAGGCGAGGAGTTTGAACACAGCTTGGCCATGATGGTGAACCCCTCTCTACTAAA	128845
QY	1071	GGTACCTGACCCCGCATGTTTGGTGAGGAGATTTCAGGAAGTCTTCTGGAAGAGGCT	1130
Db	128846	AGACAAAATTTAGCTGGGTATGTTGTCAGCGCTATATATCCAGCTACTCAGAGACT	128905
QY	1131	GTGGAGCTGAGACTCATAGATAGTGGGAGGTTTCCAGGACAGAAAGACCCAGCACC	1190
Db	128906	GAGCAGAGAAATCGCTTGATCTCTGGGAGCGAGGTTCAGTGAGCCAGATCTGTC	128965
QY	1191	TACAAAGCATGACTTTGAGAGAGCATTCATCAATTCATCACTGATGAATTTTCAGATGG	1250
Db	128966	ACCGACTCCAGCTGGGTGACAGAGTGAGACTCTCTTCAAAAAAATAAATAATTC	129025
QY	1251	GCAGCT-----GGCTCATGCTGTAATCCAGCACTTTTGGAGGCTGAATGGG	1300
Db	129026	AGGCTCACACTGTATAGCTCAGCTATATTCAGCACTTTTGGAGGCGCAGGCAGG	129085
QY	1301	AGATGACTTGGCTTGGCAATTTGTGCAAGCTTGGCAACATGTTGAGACCTTGCTC	1360
Db	129086	CAGATCACCTGAGTTCAGGCTTTCAGAGCGGCTTGGCAACATGTTGAAATGCTGCTC	129145
QY	1361	CACAAACAAACAAACAAACAAACAAATCATTTACCTGTTACCATGGTACCAG--GTAC	1418
Db	129146	TACTAAAAATACAAA--AATTAGCTGGGCTGTTGGAGGCTGCTGTAATCCAGCTGTT	129204
QY	1419	ATAGAAATGACTCAGCAGATATGTTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1478
Db	129205	GGGAGCTGAGGAGAGGATCGTTTGAACCCAGGAGGCGGGTTTACAGTAGGCGGAGA	129264
QY	1479	TGAGTAAGCAATAGAGGAGGAGGAAATATAATCTTAAATGAGAGGTACAGATTTGAG	1538
Db	129265	TTGGCCACCGACTCCAGCTGGGCGACAAAGACAGCTGTCCCTCCCAAAAAA	129324
QY	1539	CAACACAGGCGACAGGATATGACAGGTTAAAGAGGAAATCAGGAGGCTTCTCAG	1598
Db	129325	AAAGTTCACTTGACATTTGAAGTGAACCCAGTGAATCCAGGATGCTGTCA--CCTCA	129382
QY	1599	AGAAGTGCACATTTAAGCCGACATGAAGATGAACAGTGTAGTTTCAACAGGATGGGA	1658
Db	129383	GACTGAGCTTAATCATCTGCAAGATTTCTTTCCAAATTAAGGTCACTCGCATTCACA	129442
QY	1659	TGGAAGGGGTGAGAGTGAAGGAGCAGAGGAACTGCAGGATCATAGGCTTAGACAGGG	1718
Db	129443	GGTAGACATCTCATTTTGAGGAGTGAAGGAGCAAACTTCACTCTCTCTCTCTCTCTCT	129502
QY	1719	GATCTGAGCCCTTGAAGAGTGAAGAGCAGGAGTTCGTTAGTGGGTTAAGTAAC	1778
Db	129503	TCCACACTTTTGTGTAGTACAGCATCTGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	129562
QY	1779	AAAGCTGAGAGCCAGGAAATCCCTGTGTATGAGGCGCTGTGAGTCACTCAGAGTGT	1838
Db	129563	CTGGAGCCCACTTGTGGCCATTTTCAAGTTCAGAGAGTGAAGGAGTGTCTGCCCCA--	129619
QY	1839	TGGGCTTTTGTCTTCTGGGAGCAGTTCGATTTTAAAGCAGGAAACAGCTGTATTTCAG	1898
Db	129620	TAAATCAGAGCATCTCCCTCGCCATAGTATGATCCAGGAGTGGTCCAGTCAAGGCCAAT	129679
QY	1899	TTGGAGAAATCTGTGTGTGCTGAAAGGAGTGAACCTGAGGCTAGGAGCCAGGG	1958
Db	129680	CCAGGACTTCTGATGTCTTTTGCAGTAAAGGCCACTAAGAGAGCAGTGCAGCCAGCC--	129737
QY	1959	TGATAGGAGGATCCAGGCTGATGGGAGGCTGGGAGTTCGCGGTGTATGACACAGGCT	2018
Db	129738	ATCTTTCAGAGCCATCTTGTAGGCGGACCTGGCTGAGAATGAGATCCACAGAG--	129795

DEFINITION		A method of predicting cancer condition.	
ACCESSION	BD176843		
VERSION	BD176843.1	GI:29122555	
KEYWORDS	WO 02072828-A/6.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE	1 (Bases 1 to 79528)		
AUTHORS	Kato, K., Iwao, K., Noguchi, S. and Matoba, R.		
TITLE	A method of predicting cancer condition		
JOURNAL	Patent: WO 02072828-A 6 19-SEP-2002; DNA CHIP RESEARCH INC. HITACHI SOFTWARE ENGINEERING CO LTD, KIKUYA		
COMMENT	KATO, KYOKO IMAO, SHINZABURO NOGUCHI, RYO MATOBA OS Homo sapiens (human) PN WO 02072828-A/6 PD 19-SEP-2002 PF 07-MAR-2002 WO 2002JP002153 PR 14-MAR-2001 JP 01P 073063, 06-APR-2001 JP 01P 108503 PR 02-AUG-2001 JP 01P 234807 PI KIKUYA KATO, KYOKO IMAO, SHINZABURO NOGUCHI, RYO MATOBA PC C12N15/12, C12Q1/68, G06F19/00 CC A method of predicting cancer condition FH Key Location/Qualifiers FT source 1..79528 /organism='Homo sapiens (human)'		
FEATURES	source		
1..79528 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606"			
ORIGIN			
	Query Match 8.6%; Score 257; DB 6; Length 79528; Best Local Similarity 67.5%; Pred. No. 2.8e-57; Matches 410; Conservative 0; Mismatches 185; Indels 12; Gaps 3;		
QY	2407	GCTCAGCGTTGCCATTCTAGCACCTTTGGAGCGCTGAGGTGGGTGGATTACCTCAGGTCTCAG	2466
Db	64446	GCTCAGCGCTGTAAATCCAGCACCTTTGGGAGCGCGAGCGGTGGATCACCTGAGTCTCG	64505
QY	2467	GAGTTCAGACCGACGCTAGACAAAGTACAGAAACCCCATCTCTCTACTGAAGATATAAATT	2526
Db	64506	GAGTTCGAGACCGACGCTGATTAAAGTGAGAAACCCCATCTCTACTAAAAATACAAATC	64565
QY	2527	AGCTGGGCGTGTAGTGCATATGCTGTAAATCCAGCTAGTCGGGAAGCTGAGGAGGAGAA	2586
Db	64566	AGCGGGCGCATGTCACACATGCTGTAAATCCAGCTACTTTGGGAGGTGGAGGCGAGGAA	64625
QY	2587	TCGCTTGAACCGGAGAGCGAGGTTGCGGTGAGCTGAGATCGTGCCATTTCACCTCCAGC	2646
Db	64626	TCATTTGAACCGGGAAGGACAGAGGTTCGGGTGAGCCGAGATCGGCCATTGCATCTCAGC	64685
QY	2647	CTGGGTGCAAAAGCAAGACCTCGTCTCAATAATAATAATAATAATAATAATAATAATAATA	2706
Db	64686	CTGGGCAACAGAGCGAACTCCGCTCAAAAAAAGAAAAAGAAAGAAATTTGGAA	64745
QY	2707	TGGGTATCCAGCTACCTACTTTTTCAGAGAAATGTAATCTCTTACCCAGGCGCAAGGA	2766
Db	64746	TTGGTAAGTCTCTCCAACTTTGTTCTTTTTTTTTTTTCTTTTCTTTTCTTTTCTTTTCTTT	64805
QY	2767	TGGGAGAACCG- TTTGATTATGCAATTTATTGAGCACCTACTTGAGCTCTCATCTCCCTGGC	2825
Db	64806	TGGGTCTCTTGGATTGTCTTATTTGATTTTAGATGGAATTTGCAATTTCTGCAAGTAGC	64865
QY	2826	TAGGCTGGAAAT-----GGACTCAGATGGAGCTGGAAGAGTCCCTCAGGGAACCTCAC	2879
Db	64866	CAGCTGGGATTTTGAGAGGAATTCCTAGAAATCTAGAAATCTAGAAATCTAGAAATCTAG	64925
QY	2880	TAGAAAGAGGAGGAATC-----GGCGGGCGGTGGCTCAGCGCTGTAATCCCAACAC	2934
Db	64926	TCTTAACAGTATTAACTCTCCAGGCGGCGGGTGGCTCATGCTGTAATCTGTAATCCAGCA	64985

QY	2019	GGGGCCAGGGATGGGGAAGGAGTAATTTGGAGAGCGCTTGGGGCTCTTGCGCGAGGAA	2078
Db	129796	GAGAACAGGGCCGAGAGATGGAAAGTGAAGCCAGTCTTTAGTTGGGAGCCATACAC	129855
QY	2079	TGGATGGTGGGCTGAAACAGGAGAGGAGAGATGCTTAGGCCACATTTTGGAAACACAGTAGG	2138
Db	129856	CCCCTTTCTGCTGAAGCCAGCTGAGAGGGTTTCTTGGCCCTTACTCATTTGTAGTTT	129915
QY	2139	GCAAGACAGGAGACACCCAGGGGAAGTGCCCAAGAGACACAGCAGGCTGGCATTTGA	2198
Db	129916	GAAGGAAAGCAGAGGTTTAAAGGAAAAAGACACACACACACGCAACACACACACAC	129975
QY	2199	CAGGGAAGTCTGTCTGGAGCAGTCTCTTGGATAAGGG-----AGGAAATGCTGCAG	2252
Db	129976	ACACACAGAGAGAGAGAGAGTGGGGGTGGGATCGGCTCTACAGCAAACACAGGCAT	130035
QY	2253	TTCCATCTCTCTCCCTCTCTGTTCAACCTTAAACTACATGGGGACAGGACCCAGTGGG	2312
Db	130036	TTATGTTCCAGCATCAGCCCTGCGAGGTGGGACCAAGCTTGATGCCAGTGGCTACTGCCA	130095
QY	2313	ACTCCATAAATGATGGATGGGTGGATGAGAGAAAGGAGGAGGAACAACTCTTCATTC	2372
Db	130096	CTTATAGGCTTGGAGTACGTATAGGCCCTGGGCAAGGAGGGTCTGGGGGGCATGGCCTGCTG	130155
QY	2373	ATCTGCTGTTATTACAGAACAGCGCAGTTCGCGTG-CTCAGCGTTGCCATTCTAGCACATT	2431
Db	130156	CCTGGAAAAATGTTTCATAGCAGCTGGGCGCGTGGCTCACACCTGTATCCAGCACATT	130215
QY	2432	TGGGAGGCTCAGGTGGGTGGATTACCTCAGGTCAAGGATTAAGACAGCAGCTAGACACG	2491
Db	130216	TAGGAGCCGAGGTGGGTGCATCAC--AAGGTCAGGAGTTCCAGACAGCAGCTGGCCAAAC	130273
QY	2492	TAGAGAAACCCATCTCTACTGAAGATA-TAAAAATTAGCTGGCGTAGTGGCATATGCTCT	2550
Db	130274	TAGTGAACCCCTGTCTCTGCTTAAATAATACAAAAATTAGCTGGGCATGGTGGCAGGCACCT	130333
QY	2551	GTAATCCAGCTAGTTCGGGAAGCTGAGGCAAGGAGAAATCGCTTGAACCCGAGAGGCGAGG	2610
Db	130334	GTAATCTTAGCTACTCAGAGGCTGAGACAGGAGNATCACTTGAACCTGGAGGTAGATG	130393
QY	2611	TTGCGGTGAGCTGAGATGGTGCCATTGCACTCGAGCCTGGGTGACAAAGCAAGACCTCGT	2670
Db	130394	TTGCAGTCAGCCGAGATAGCACCACTGCATCTCCATCTCTGGGCAACAAAGCAAGACTCTGT	130453
QY	2671	CTCAATAATAATAATTAACAAACAGAGAGGAGCTGGGTGATCCAGCTACTACTTT	2730
Db	130454	CTCAAAAAACAAAAAGAAAGTTTAATACATGCCATGTCAGGCAGTTTGGCCCTTGT	130513
QY	2731	TCAGGAGAAATGTAC---TCCCTTACCACAGGCGAAAGGATGGGAGAACCAAGTTTGATTAT	2787
Db	130514	CCACAGATGTGACAGTATGCTTTCTCAGCTGGGTCTTTGCTGGCAGGATATGATAAG	130573
QY	2788	GCATTTATGAGCACTTACTGAGTCTCTATCCCTGGGCTAGGCTGGAAATGGAATCTCAGATG	2847
Db	130574	AAAGTCAGCGAGTCAGGAGGAGATGTTTGTACACCTGAAACCCCTGTGGAAATTTTCACT	130633
QY	2848	GAGCTGAGAGTCCCCCTCAGGAAACCTCACTAGAAAGAGGAGGAATCGGCGGGCGC	2907
Db	130634	TTGACCAAGGCTGTGAATATGCGGGGGCTTACAAATGGTGAGTTTGGCGGGCAC	130693
QY	2908	GGTGGCTCACGCTGTAAATCCCAACACTTTGGGAGGTGAGGTGGGTGGATACAAAGTTC	2967
Db	130694	GGTGGCTCACACCTGTAATCCCAACCACTTTGGGAGCGCGAGCGGAGCAGATCAC-CTGAG	130752
QY	2968	AGGAGATCGAGACCACTCTGGCTAACACAGTGAA 3001	
Db	130753	GTGAGTTCGGGACCGCTGGCCAAACATGGTGAA 130786	

RESULT 15



Qy	2935	TTTGGGAGGCTGAGGTGGTGGATCACAAGGTCAGGAGATCGAGACCATCCTGGCTAACA	2994
Db	64986	TTTTGGAGGTGAAGGTGGACGGATCACCGAGGTCAGGAGATCGAGACCATCCTGGCTAACA	65045
Qy	2995	CAGTGAA	3001
Db	65046	CGGTGAA	65052

Search completed: September 17, 2004, 14:58:13  
Job time : 11445.9 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 17, 2004, 06:15:30 ; Search time 1107.52 Seconds  
(without alignments)  
11511.184 Million cell updates/sec

Title: US-10-669-693-3\_COPY\_1980\_4980

Perfect score: 3001

Sequence: 1 cttcttcctccctaccctccccc.....atcctgggtaacacagtgaa 3001

Scoring table: IDENTITY\_NUC

Gapop 10.0 ; Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N Geneseq\_29Jan04:\*  
1: Geneseqn1980s:\*  
2: Geneseqn1990s:\*  
3: Geneseqn2000s:\*  
4: Geneseqn2001as:\*  
5: Geneseqn2001bs:\*  
6: Geneseqn2002s:\*  
7: Geneseqn2003as:\*  
8: Geneseqn2003bs:\*  
9: Geneseqn2003cs:\*  
10: Geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	3001	100.0	17752	6	AAD45439 Human dru
2	257	8.6	79528	6	AAL50814 Human can
3	253	8.4	16747	4	AAL90152 Human dig
4	253	8.4	16747	4	AAL80326 Human imm
5	253	8.4	16747	4	AAL86604 Human imm
6	253	8.4	16747	4	AAL36989 Human mus
7	253	8.4	16747	7	ABX59977 CDNA enco
8	247.6	8.3	110096	6	ABN95044 Gene #154
9	247.2	8.2	2981	5	ABAL7151 Human ner
10	247.2	8.2	5041	5	ABAL7153 Human ner
11	245.8	8.2	110000	6	ABQ74964 5
12	245.6	8.2	96593	9	Continuation (6 of
13	245.6	8.2	96594	8	ADC85367 Mouse BLM
14	245.6	8.2	96594	9	ADA02888 Human BLM
15	245.6	8.2	99500	6	ADB72626 Human BLM
16	244.8	8.2	10500	6	ABK13086 Human REC
17	244.8	8.2	23130	4	AAD41740 Human enco
18	242	8.1	5952	9	AAL77045 Human imm
19	241.8	8.1	154465	6	ADB54002 PCNA geno
20	241.8	8.1	158245	6	AAD28763 Human AKA
21	241.8	8.1	161425	4	AAD28762 Human AKA
22	241.8	8.1	162025	4	AAL02340 Human AKA
23	241.8	8.1	162025	6	AAL02339 Human AKA
					AAD28758 Human AKA

24	241.8	8.1	162025	6	AAD28759	Human AKA
25	240.8	8.0	17327	2	AAQ44278	Serglycin
26	240.6	8.0	185371	6	ABT10718	Human bre
27	240.2	8.0	98829	8	AAD55116	Human ACE
28	240	8.0	154902	6	ABQ88198	Human ost
29	238.8	8.0	8624	6	AAD44437	Human coa
30	238.8	8.0	10953	5	AAS29204	Genomic s
31	238.8	8.0	10953	6	ABS68344	Human DNA
32	238.8	8.0	10953	9	ADC25466	Human CDN
33	238.8	8.0	22645	4	AAK89727	Human dig
34	238.8	8.0	22645	4	AAK105495	Human rep
35	238.8	8.0	22645	4	AAK04985	Human rep
36	238.8	8.0	22645	4	ABL97878	Human tes
37	238.8	8.0	22645	4	ABL97878	Human tes
38	238.8	8.0	113604	8	ADA42607	Human IQT
39	238.8	8.0	113604	8	ADA42606	Human IQT
40	238	7.9	357	4	AAK76716	Human imm
41	238	7.9	14874	4	AAK89726	Human dig
42	238	7.9	14874	4	AAL04984	Human dig
43	238	7.9	14874	4	ABL97877	Human rep
44	237.8	7.9	12452	4	AAL37624	Human mus
45	237.8	7.9	12452	5	ABA21062	Human ner

## ALIGNMENTS

## RESULT 1

AAD45439

ID AAD45439 standard; DNA; 17752 BP.

XX

AC AAD45439;

XX

DT 27-DEC-2002 (first entry)

XX Human drug-metabolising enzyme encoding gene.

DE Human; drug-metabolising enzyme; cytochrome p450; therapy; cytostatic;

KW Parkinson's disease; myotonic dystrophy; developmental defect; enzyme;

KW cancer; chromosome 19; gene; ds.

XX Homo sapiens.

OS

XX

PH Key

FT CDS

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT









```

PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225575P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0232196P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 23-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0251989P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 11-DEC-2000; 2000US-0251990P.
PR 05-JAN-2001; 2001US-0259678P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-502630/55.
XX
XX Polynucleotides encoding digestive system antigens, useful for
XX diagnosing, treating, preventing and/or prognosing disorders of the
XX digestive system, particularly cancer and cancer metastases.
XX
XX Disclosure; SEQ ID NO 3728; 986pp; English.
XX
XX The present invention provides the protein and coding sequences of a
XX number of human digestive system antigens. These can be used in the
XX diagnosis, treatment and prevention of digestive system disorders,
XX including cancer, Meckel's diverticulum, bacterial or parasitic
XX infections, appendicitis, Hirschsprung's disease, chronic colitis or
XX ulcerative colitis. The present sequence is a genomic DNA fragment
XX encoding a digestive system antigen of the invention
XX
XX SQ Sequence 16747 BP; 4203 A; 3717 C; 3840 G; 3787 T; 0 U; 1200 Other;
Query Match 8.4%; Score 253; DB 4; Length 16747;
Best Local Similarity 27.2%; Pred. No. 1.5e-55;
Matches 659; Conservative 0; Mismatches 1740; Indels 22; Gaps 6;
QY 301 ACAGTCCCATCAGCGGGTGAGTGGCTGAGTCCGAAATCCCAACACATTTGGAGGCG 360
Db 16152 AAAATGAGTTTGGCTGGCGGGTGCTCAGGCGCTATATATCCCAACATTTGGAGGCC 16093
QY 361 GAGGTGGTGGATCATCTTGGAGTTCAGAGTTCAGACCGCTGGCCACATCGTGAAC 420

```



QY 2560 GCTAGTCGGGAAGCTGAGGACGAGAGATCGCTTGAAACCGAGAGGCACAGGTTGCGGTGA 2619  
|||||  
Db 13873 GCTACTTAGGAGGCTGAGGAGGAGATTGCTTGAGTCCGCGCAGGTCAAGGCTGCAGTGA 13814  
|||||  
QY 2620 GCTGAGATCGTGCATTCGACTCCAGCCTGGGTGACAAAGCAAGACCTCGTCTCAATAAT 2679  
|||||  
Db 13813 ACAATGATCAAGCCACTGCACTCCAGCTTGAGCGACAGTGAACACTGCTCTCAATAAAA 13754  
|||||  
QY 2680 AATAATTAATACAAAACAGAA 2700  
|||||  
Db 13753 CCCAAAAACCAACAAAAACAA 13733  
|||||  
RESULT 4  
AAK80326/C  
ID AAK80326 standard; DNA; 16747 BP.  
XX  
AC AAK80326;  
XX  
XX  
DT 07-NOV-2001 (first entry)  
XX  
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:35138.  
XX  
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;  
KW cytostatic; gene therapy; vaccine; metastasis; ds  
XX  
OS Homo sapiens.  
XX  
PN WO200157182-A2.  
XX  
XX  
PD 09-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US001354.  
XX  
PR 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214866P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 14-JUL-2000; 2000US-0217496P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226686P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.





PR	08-SEP-2000;	2000US-0231243P;
PR	08-SEP-2000;	2000US-0231244P;
PR	08-SEP-2000;	2000US-02312413P;
PR	08-SEP-2000;	2000US-0231141P;
PR	08-SEP-2000;	2000US-0231080P;
PR	08-SEP-2000;	2000US-0232081P;
PR	14-SEP-2000;	2000US-0231968P;
PR	14-SEP-2000;	2000US-023397P;
PR	14-SEP-2000;	2000US-023398P;
PR	14-SEP-2000;	2000US-023399P;
PR	14-SEP-2000;	2000US-0232400P;
PR	14-SEP-2000;	2000US-0232401P;
PR	14-SEP-2000;	2000US-0233063P;
PR	14-SEP-2000;	2000US-0233064P;
PR	14-SEP-2000;	2000US-0233065P;
PR	14-SEP-2000;	2000US-0234223P;
PR	21-SEP-2000;	2000US-0234274P;
PR	21-SEP-2000;	2000US-0234997P;
PR	25-SEP-2000;	2000US-0234998P;
PR	25-SEP-2000;	2000US-0234999P;
PR	26-SEP-2000;	2000US-0235484P;
PR	27-SEP-2000;	2000US-0235834P;
PR	27-SEP-2000;	2000US-0235836P;
PR	29-SEP-2000;	2000US-0236327P;
PR	29-SEP-2000;	2000US-0236367P;
PR	29-SEP-2000;	2000US-0236368P;
PR	29-SEP-2000;	2000US-0236369P;
PR	29-SEP-2000;	2000US-0236370P;
PR	29-SEP-2000;	2000US-0236370P;
PR	02-OCT-2000;	2000US-0236802P;
PR	02-OCT-2000;	2000US-0237037P;
PR	02-OCT-2000;	2000US-0237038P;
PR	02-OCT-2000;	2000US-0237039P;
PR	02-OCT-2000;	2000US-0237040P;
PR	13-OCT-2000;	2000US-0237935P;
PR	13-OCT-2000;	2000US-0239327P;
PR	13-OCT-2000;	2000US-0240560P;
PR	20-OCT-2000;	2000US-0241221P;
PR	20-OCT-2000;	2000US-0241628P;
PR	01-NOV-2000;	2000US-0244617P;
PR	08-NOV-2000;	2000US-0244747P;
PR	08-NOV-2000;	2000US-0246475P;
PR	08-NOV-2000;	2000US-0246476P;
PR	08-NOV-2000;	2000US-0246477P;
PR	08-NOV-2000;	2000US-0246478P;
PR	08-NOV-2000;	2000US-0245523P;
PR	08-NOV-2000;	2000US-0245524P;
PR	08-NOV-2000;	2000US-0245525P;
PR	08-NOV-2000;	2000US-0245526P;
PR	08-NOV-2000;	2000US-0245527P;
PR	08-NOV-2000;	2000US-0245528P;
PR	08-NOV-2000;	2000US-0245532P;
PR	08-NOV-2000;	2000US-0246609P;
PR	08-NOV-2000;	2000US-0246610P;
PR	08-NOV-2000;	2000US-0246611P;
PR	08-NOV-2000;	2000US-0246613P;
PR	17-NOV-2000;	2000US-0249207P;
PR	17-NOV-2000;	2000US-0249208P;
PR	17-NOV-2000;	2000US-0249209P;
PR	17-NOV-2000;	2000US-0249210P;
PR	17-NOV-2000;	2000US-0249211P;
PR	17-NOV-2000;	2000US-0249212P;
PR	17-NOV-2000;	2000US-0249213P;
PR	17-NOV-2000;	2000US-0249214P;
PR	17-NOV-2000;	2000US-0249215P;
PR	17-NOV-2000;	2000US-0249216P;
PR	17-NOV-2000;	2000US-0249217P;
PR	17-NOV-2000;	2000US-0249218P;
PR	17-NOV-2000;	2000US-0249244P;
PR	17-NOV-2000;	2000US-0249245P;

PR	17-NOV-2000; 2000US-0249264P.	
PR	17-NOV-2000; 2000US-0249265P.	
PR	17-NOV-2000; 2000US-0249297P.	
PR	17-NOV-2000; 2000US-0249299P.	
PR	17-NOV-2000; 2000US-0249300P.	
PR	01-DEC-2000; 2000US-0250160P.	
PR	01-DEC-2000; 2000US-0250311P.	
PR	05-DEC-2000; 2000US-0251030P.	
PR	05-DEC-2000; 2000US-0251988P.	
PR	05-DEC-2000; 2000US-0256719P.	
PR	06-DEC-2000; 2000US-0251479P.	
PR	08-DEC-2000; 2000US-0251856P.	
PR	08-DEC-2000; 2000US-0251868P.	
PR	08-DEC-2000; 2000US-0251869P.	
PR	08-DEC-2000; 2000US-0251989P.	
PR	08-DEC-2000; 2000US-0251990P.	
PR	11-DEC-2000; 2000US-0254097P.	
PR	05-JAN-2001; 2001US-0259678P.	
XX		
PA	(HUMA-) HUMAN GENOME SCI INC.	
XX		
PI	Rosen CA, Barash SC, Ruben SM;	
XX		
DR	WPI; 2001-483426/52.	
XX		
PT	Nucleic acids encoding human immune/hematopoietic antigen polypeptides,	
PT	useful for preventing, diagnosing and/or treating cancers and metastasis.	
XX		
PS	Disclosure; SEQ ID NO 41416; 3071pp + Sequence Listing; English.	
XX		
CC	AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)	
CC	amino acid sequences given in AAM82170 to AAM91921. (I) have cytosstatic	
CC	activity, and can be used in gene therapy and vaccine production. (I)	
CC	proteins and polynucleotides may be used in the prevention, diagnosis and	
CC	treatment of diseases associated with inappropriate (I) expression. For	
CC	example, they may be used to treat disorders associated with decreased	
CC	expression by rectifying mutations or deletions in a patient's genome	
CC	that affect the activity of (I) by expressing inactive proteins or to	
CC	supplement the patients own production of (I). Additionally, (I)	
CC	polynucleotides may be used to produce the secreted (I), by inserting the	
CC	nucleic acids into a host cell and culturing the cell to express the	
CC	protein. (I) proteins and polynucleotides may be used to prevent,	
CC	diagnose and treat immune/haematopoietic-related diseases, especially	
CC	cancers and cancer metastases of haematopoietic-derived cells. AAK64703	
CC	to AAK87694 represent human immune/haematopoietic antigen genomic	
CC	sequences from the present invention. AAK54942 to AAK54950 and AAM82169	
CC	represent sequences used in the exemplification of the present invention	
XX		
SQ	Sequence 16747 BP; 4203 A; 3717 C; 3840 G; 3787 T; 0 U; 1200 Other;	
	Query Match 8.4%; Score 253; DB 4; Length 16747;	
	Best Local Similarity 27.2%; Pred. No. 1.5e-55;	
	Matches 659; Conservative 0; Mismatches 1740; Indels 22; Gaps 6;	
QY	301 ACAGTCCCATCAGCGGGTGCAGTGGCTAGTGCCTGAAATCCCAACACTTTGGGAGGCG 360	
Db	16152 AAAATAGATTTAGCTGGGCGCGTGCTCAGGCTTATATCCCAACACTTTGGGAGGCC 16093	
QY	361 GAGGTGGTGGATCACTTTGAGGTCCAGAGTTCCAGACCAAGCTGGCCAAATGGTGAAC 420	
Db	16092 GAGGCTGGTGGATCCACTTTGAGGTCCAGAGTTAGAGACCAAGCTGGCCAAATGGTGAAC 16033	
QY	421 CTTCTCTCTACTAAATA-TAAAAATAGCTGGGATGGTGGTACCTGTAATCCC 479	
Db	16032 TCTGTCTCTACTAAATTTACAAAATATAGTGGGCATGGTGGCGGTGCCCTGTATCCC 15973	
QY	480 AGATACTTGGGAGGTTGAGGCAGGAGAATCGCTTTGAACCCGGAGGCAGATGTTGCAGTG 539	
Db	15972 AGCTACTTGGGAGGCTGAGGCAGAGAATTTGCCCTGAACCAAGAGATGGAGGTGTAGTG 15913	
QY	540 AACCAAGACTGTGCCACTGCACCTCCAGTCTGGGCAACA-GAGTGAGCCTCCATCTCAAC 598	
Db	15912 AGCCAAAGCTGCGCACTGCACCTAGCTTGGGCAACAGAGGGAACTCCGCTCAAAA 15853	





RESULT 6	PR	08-SEP-2000;	2000US-0231414P.
AAL36989/c	PR	08-SEP-2000;	2000US-0232080P.
ID AAL36989 standard; DNA; 16747 BP.	PR	08-SEP-2000;	2000US-0232081P.
XX	PR	12-SEP-2000;	2000US-0231968P.
AC AAL36989;	PR	14-SEP-2000;	2000US-0232397P.
XX	PR	14-SEP-2000;	2000US-0232398P.
DT 08-JAN-2002 (first entry)	PR	14-SEP-2000;	2000US-0232399P.
XX	PR	14-SEP-2000;	2000US-0232400P.
DE Human musculoskeletal system related polynucleotide SEQ ID NO 3354.	PR	14-SEP-2000;	2000US-0232401P.
XX	PR	14-SEP-2000;	2000US-0233063P.
XX	PR	14-SEP-2000;	2000US-0233064P.
KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;	PR	14-SEP-2000;	2000US-0233065P.
KW anti-allergic; hepatotropic; antidiabetic; antiinflammatory; antitumor;	PR	21-SEP-2000;	2000US-0234223P.
KW vulnery; aniconvulsant; antibacterial; antifungal; antiparasitic;	PR	21-SEP-2000;	2000US-0234274P.
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;	PR	25-SEP-2000;	2000US-0234997P.
KW neurological disease; infection; human; secreted protein;	PR	25-SEP-2000;	2000US-0234998P.
KW musculoskeletal system; ds.	PR	26-SEP-2000;	2000US-0235484P.
XX	PR	27-SEP-2000;	2000US-0235834P.
OS Homo sapiens.	PR	27-SEP-2000;	2000US-0235836P.
XX	PR	29-SEP-2000;	2000US-0236327P.
XX	PR	29-SEP-2000;	2000US-0236367P.
PN WO200155367-A1.	PR	29-SEP-2000;	2000US-0236368P.
XX	PR	29-SEP-2000;	2000US-0236369P.
PD 02-AUG-2001.	PR	29-SEP-2000;	2000US-0236370P.
XX	PR	02-OCT-2000;	2000US-0236802P.
PF 17-JAN-2001; 2001WO-US001338.	PR	02-OCT-2000;	2000US-0237037P.
XX	PR	02-OCT-2000;	2000US-0237038P.
XX	PR	02-OCT-2000;	2000US-0237039P.
PR 31-JAN-2000; 2000US-0179065P.	PR	02-OCT-2000;	2000US-0237040P.
PR 04-FEB-2000; 2000US-0180628P.	PR	02-OCT-2000;	2000US-0239935P.
PR 24-FEB-2000; 2000US-0184664P.	PR	13-OCT-2000;	2000US-0239937P.
PR 02-MAR-2000; 2000US-0186350P.	PR	20-OCT-2000;	2000US-0240960P.
PR 16-MAR-2000; 2000US-0189874P.	PR	20-OCT-2000;	2000US-0241221P.
PR 17-MAR-2000; 2000US-0190076P.	PR	20-OCT-2000;	2000US-0241785P.
PR 18-APR-2000; 2000US-0198123P.	PR	20-OCT-2000;	2000US-0241786P.
PR 19-MAY-2000; 2000US-0205515P.	PR	20-OCT-2000;	2000US-0241787P.
PR 07-JUN-2000; 2000US-0209467P.	PR	20-OCT-2000;	2000US-0241808P.
PR 28-JUN-2000; 2000US-0215135P.	PR	20-OCT-2000;	2000US-0241809P.
PR 30-JUN-2000; 2000US-0216647P.	PR	20-OCT-2000;	2000US-0241826P.
PR 07-JUL-2000; 2000US-0216880P.	PR	01-NOV-2000;	2000US-0244617P.
PR 11-JUL-2000; 2000US-0217487P.	PR	08-NOV-2000;	2000US-0246474P.
PR 11-JUL-2000; 2000US-0217496P.	PR	08-NOV-2000;	2000US-0246475P.
PR 14-JUL-2000; 2000US-0218290P.	PR	08-NOV-2000;	2000US-0246476P.
PR 26-JUL-2000; 2000US-0220963P.	PR	08-NOV-2000;	2000US-0246477P.
PR 26-JUL-2000; 2000US-0220964P.	PR	08-NOV-2000;	2000US-0246478P.
PR 14-AUG-2000; 2000US-0224518P.	PR	08-NOV-2000;	2000US-0246523P.
PR 14-AUG-2000; 2000US-0224519P.	PR	08-NOV-2000;	2000US-0246524P.
PR 14-AUG-2000; 2000US-0225213P.	PR	08-NOV-2000;	2000US-0246525P.
PR 14-AUG-2000; 2000US-0225214P.	PR	08-NOV-2000;	2000US-0246526P.
PR 14-AUG-2000; 2000US-0225267P.	PR	08-NOV-2000;	2000US-0246527P.
PR 14-AUG-2000; 2000US-0225268P.	PR	08-NOV-2000;	2000US-0246528P.
PR 14-AUG-2000; 2000US-0225270P.	PR	08-NOV-2000;	2000US-0246532P.
PR 14-AUG-2000; 2000US-0225447P.	PR	08-NOV-2000;	2000US-0246609P.
PR 14-AUG-2000; 2000US-0225575P.	PR	08-NOV-2000;	2000US-0246610P.
PR 14-AUG-2000; 2000US-0225758P.	PR	08-NOV-2000;	2000US-0246611P.
PR 14-AUG-2000; 2000US-0225759P.	PR	08-NOV-2000;	2000US-0246613P.
PR 18-AUG-2000; 2000US-0226279P.	PR	17-NOV-2000;	2000US-0249207P.
PR 22-AUG-2000; 2000US-0226581P.	PR	17-NOV-2000;	2000US-0249208P.
PR 22-AUG-2000; 2000US-0226868P.	PR	17-NOV-2000;	2000US-0249209P.
PR 22-AUG-2000; 2000US-0227182P.	PR	17-NOV-2000;	2000US-0249210P.
PR 23-AUG-2000; 2000US-0227009P.	PR	17-NOV-2000;	2000US-0249211P.
PR 30-AUG-2000; 2000US-0228924P.	PR	17-NOV-2000;	2000US-0249212P.
PR 01-SEP-2000; 2000US-0229387P.	PR	17-NOV-2000;	2000US-0249213P.
PR 01-SEP-2000; 2000US-0229343P.	PR	17-NOV-2000;	2000US-0249214P.
PR 01-SEP-2000; 2000US-0229344P.	PR	17-NOV-2000;	2000US-0249215P.
PR 01-SEP-2000; 2000US-0229345P.	PR	17-NOV-2000;	2000US-0249216P.
PR 05-SEP-2000; 2000US-0229509P.	PR	17-NOV-2000;	2000US-0249217P.
PR 05-SEP-2000; 2000US-0229513P.	PR	17-NOV-2000;	2000US-0249218P.
PR 06-SEP-2000; 2000US-0230437P.	PR	17-NOV-2000;	2000US-0249244P.
PR 06-SEP-2000; 2000US-0230438P.	PR	17-NOV-2000;	2000US-0249245P.
PR 08-SEP-2000; 2000US-0231242P.	PR	17-NOV-2000;	2000US-0249264P.
PR 08-SEP-2000; 2000US-0231243P.	PR	17-NOV-2000;	2000US-0249265P.
PR 08-SEP-2000; 2000US-0231244P.	PR	17-NOV-2000;	2000US-0249266P.
PR 08-SEP-2000; 2000US-0231413P.	PR	17-NOV-2000;	2000US-0249297P.

17-NOV-2000; 2000US-0249299P.  
 PR 17-NOV-2000; 2000US-0249300P.  
 PR 01-DEC-2000; 2000US-0250160P.  
 PR 01-DEC-2000; 2000US-0250391P.  
 PR 05-DEC-2000; 2000US-0251030P.  
 PR 05-DEC-2000; 2000US-0251988P.  
 PR 06-DEC-2000; 2000US-0256719P.  
 PR 06-DEC-2000; 2000US-0251479P.  
 PR 08-DEC-2000; 2000US-0251856P.  
 PR 08-DEC-2000; 2000US-0251868P.  
 PR 08-DEC-2000; 2000US-0251869P.  
 PR 08-DEC-2000; 2000US-0251989P.  
 PR 08-DEC-2000; 2000US-0251990P.  
 PR 11-DEC-2000; 2000US-0254097P.  
 PR 05-JAN-2001; 2001US-0259678P.  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 PA Rosen CA, Barash SC, Ruben SM;  
 XX WPI; 2001-451937/48.  
 DR  
 XX  
 PT Isolated polypeptide for treating, preventing and/ or prognosing  
 PT disorders related to the musculoskeletal system including musculoskeletal  
 PT cancers and also for testing and detection e.g. diagnosis.  
 XX  
 PS Example 2; SEQ ID NO 3354; 781pp + Sequence Listing; English.  
 XX  
 CC The invention relates to novel genes (AAL34669-AAL37666) and proteins  
 CC (AB03087-AB04109) associated with the musculoskeletal system useful for  
 CC preventing, treating or ameliorating medical conditions e.g. by protein  
 CC or gene therapy. The genes are isolated from a range of human tissues  
 CC disclosed in the specification. The nucleic acids, proteins, antibodies  
 CC and (ant)agonists are useful in the diagnosis, treatment and prevention  
 CC of: (a) cancer, e.g. breast and ovarian cancer and other cancers of the  
 CC adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver,  
 CC lung, or urogenital; (b) immune disorders e.g. Addison's disease,  
 CC allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,  
 CC diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid  
 CC arthritis and ulcerative colitis; (c) cardiovascular disorders such as  
 CC myocardial ischaemia; (d) wound healing; (e) neurological diseases e.g.  
 CC cerebral anoxia and epilepsy; and (f) infectious diseases such as viral,  
 CC bacterial, fungal and parasitic infections. Note: The sequence data for  
 CC this patent did not form part of the printed specification, but was  
 CC obtained in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 16747 BP; 4203 A; 3717 C; 3840 G; 3787 T; 0 U; 1200 Other;  
 Query Match 8.4%; Score 253; DB 4; Length 16747;  
 Best Local Similarity 27.2%; Pred. No. 1.5e-55;  
 Matches 659; Conservative 0; Mismatches 1740; Indels 22; Gaps 6;  
 QY 301 ACAGTCCCATCAGCGGGTGCAGTGTGCTAGTGCCTGAAATCCCAACATTTGGAGGCG 360  
 DB 16152 AAAATGAGTTTAGGCTGGGCGGTGGCTCAGGCCTATATCCCAACATTTGGAGGCC 16093  
 QY 361 GAGGTGGGTGCATCAGTGTGAGGTGAGGTCGAGACACCGCTGGCCAAATGGTGAAC 420  
 DB 16092 GAGGCTGGTGGATCAGTGTGAGGTGAGGTTAGAGACCGCTGGCCAAATGGTGAAC 16033  
 QY 421 CTTGCTCTACTAAAAATA-TAAAAATTAGCTGGGCATGTGTGGTCTCTGTAATCCC 479  
 DB 16032 TCTGCTCTACTAAATTACAAAAATTAGCTGGGCATGTGTGGGCTCTGTAATCCC 15973  
 QY 480 AGATCTGGGAGGTTGAGCAGAGAAATCGTTGAACCGCGGAGGAGAGATTTGCAGTG 539  
 DB 15972 AGCTACTGGAGGCTGAGCAGAGAAATTCCTGAAACCAAGAGATGGAGTTGTAGT 15913  
 QY 540 AACCAAGATGTGCCACTGCAGTCTGGGCAACA-CAGTGAGCCTCCATCTCAAAC 598  
 DB 15912 AGCCAAGATGCGCCACTGCACTCTAGCCTGGGCAACACAGCGGAACTCCGCTCAAAA 15853

QY 599 AAACAAACAAAGAGCAGTCCCATCATGTAGGATTGAGTGTAGTGTAGTGTAGGACTGAGCCT 658  
 DB 15852 AAAAAAATAAATTTGGAGGGAGATGAGATTGTAATGTAAATATATAGACATCAATTTTA 15793  
 QY 659 TGTGCAAAAGTGAAGCACTCAATCAACAGGTTGTAGTATCAGTGTATACCATCATGAT 718  
 DB 15792 TAGGAATAGTGTAGTCTCTGAAGACCTCTATGATGGTTAATTTGGGAGATGTTTCA 15733  
 QY 719 CCAGTAAAGCCCTGAGGGTTTCAAGAGATGCCGAGCGCTTTCAAGTGTCTGGGATTTG 778  
 DB 15732 TCTGAGTCAGAAAAATATTTACAGACAATTTTGTCTGTCAGTTCAGAGGATGGAGACAGA 15673  
 QY 779 GTGGCAAGCCCTCGAATAATAGAAACAGTTCTCTGTATTATACAAAGAGAGAGAGGCC 838  
 DB 15672 ATGGAGGAGAGAGAGGCAATTCAGAGAGTGGTGGCAATTTATTTCAAGTGAGAGTAAGACAG 15613  
 QY 839 CATGCTGGTGTCTGCCAGGAATCAGTAGTAAGACAGCACCGGTCTGCTCCCA 898  
 DB 15612 TATCAGTAGAGTGAAGAGAGTAAAGACACATGCCAAAGATTAAATCAAGGATAAATA 15553  
 QY 899 GCGCACCTAGCCAGTGGGAAACAGACTCCACACAGTCCCGAGCCAGAGTGTCTCAGG 958  
 DB 15552 TTGAAGGTAATTTGATAGTTTCTGCTCGGGCAGCCACTAGTTAAGTAATATAGTTGACA 15493  
 QY 959 GCCAAGATGGGAGACAGCGGGAG-----AAAGTCAAGGTGGGATGGGAGAGGG 1008  
 DB 15492 GCGAGTTTGGAGATAAGATCCTGTTGAATTTCTTGTGTTAAGGCTTTCTGAGGATTC 15433  
 QY 1009 TCAGGCAAGAGGGGTTCAGGCCAGGCTGAGGGAAGCCCTGGGACTGTAGGAATTTAGAG 1068  
 DB 15432 TCAGTGGAATTTGTAGGTTATAGTTTGGGTGGAAATTTCTGGTTTACCAATTCGTG-G 15374  
 QY 1069 GAGTACTGACCCCGCATGTTTGGTGGAGGAGATTCAGGAAGTCTTCTCTGGAAGAGAGG 1128  
 DB 15373 TAGCTGAAGCAAGCAACAAGAAATATCTTGATCAAGGGTATGTAAAGTCATGAGGAAGG 15314  
 QY 1129 CTGTGCGAGTGCAGTCAATAGATGATGGGGAGGTGTTC--AGGCAAGAAAGACCA 1185  
 DB 15313 AAAAAGAGCTTAAATGAGAAAGGGTGTATAAAGAGAGACAGATGGAAGAGCATT 15254  
 QY 1186 GCACCTACAAAAGCATGACTTTGAGAGAGCATTCATCCATTCATGATCAATTTTCAG 1245  
 DB 15253 TTTTCTTTTCTTTTCTGNN 15194  
 QY 1246 ACTGGGCACTGCTGCTCATGCTGTAATCCAGCATTTTGAAGAGCTGAATGGGAGGAT 1305  
 DB 15193 NNN 15134  
 QY 1306 GACTTGAGCCTGAGCATTTGTGCAAGCCTGGGCAACATGGTGAGAGCCCTGCCTCCCAA 1365  
 DB 15133 NNN 15074  
 QY 1366 AACAAACAAACAAACAAATAATCAATCTGCTGCTACCATGGGTACAGGTACATAGAAA 1425  
 DB 15073 NNN 15014  
 QY 1426 TGACTCAGGAGATATGGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1485  
 DB 15013 NNN 14954  
 QY 1486 AGACAATAGAGGAGGAATAATACTTAATATGAGAGGTACAGATTGTAGAGCAACAC 1545  
 DB 14953 NNN 14894  
 QY 1546 AGGGCACAGCATATGTCAGAGGTAAAGAGGAAATCAGGGAAGGCTTCTCAGAGAGGT 1605  
 DB 14893 NNN 14834  
 QY 1606 GACATTTAAGCCCGGACATGAAGGATGAACGAGTTAGTTTCAACCAAGATGGGATGAAAG 1665  
 DB 14833 NNN 14774  
 QY 1666 GGGTGAGTGTGGAGGAGGAGGAACTGCAGGATCATAGGCCTAGACAGGGGATCTCTG 1725

[illegible]





QY 1486 AGACAATAGAGGGAGGAATATACTCTAAATGAGAGGTACAGATTGTGAGACAAACAC 1545  
Db 14953 NNN 14894  
QY 1546 AGGCGACAGGCATATGTACGAGGGTAAAGAGGGAATCAGGGAAGGCTTCTCAGAAAGGT 1605  
Db 14893 NNN 14834  
QY 1606 GACATTTAAGCCGGGACATGAAGGATGAACGAGTTAGTTCCACCAAGGATGGGATGGAAG 1665  
Db 14833 NNN 14774  
QY 1666 GGGTCAGAGTGATGAGGCGAGAGGGAACCTGCAGGATCATAGCCCTAGACAGGGGATCCTG 1725  
Db 14773 NNN 14714  
QY 1726 ACGCCCTTGAGGAAGTGAGAGAAGACCAGCGCAGTCGTAGTGGGTAAAGTAAACAAAGCTG 1785  
Db 14713 NNN 14654  
QY 1786 AGAAGCCAGGGAATCCCTGGTCATGACGGGCTGTGAGTCAGTCAGAGTGTTGGGCT 1845  
Db 14653 NNN 14594  
QY 1846 TTTGTTTTCCTGGGAGCAGTCGATTTTAAGCAGGGAACAGCTGTATTTCAGAGTTGGGAA 1905  
Db 14593 NNN 14534  
QY 1906 GATCCTGTGTTGCTGCTGAAGGGGATGAACCTGGAGGCTAGGAGCCAGGCTGATAGG 1965  
Db 14533 NNN 14474  
QY 1966 GAGGATCCAGGGTGATGGGGAGGCTGGGAGGTCGCGGTGATGGACAGGGCTGGGGCCA 2025  
Db 14473 NNN 14414  
QY 2026 GGGGATGGGAGAGGAGTAAITGGGAGAGGCTGGGCTCTGGCCGAGGAATGGATGG 2085  
Db 14413 NNN 14354  
QY 2086 TGGGCTGAACAGGAGAGGAGAGATGCTTAGGCCACTTTGGACACAGTAGGGCAAGGA 2145  
Db 14353 NNN 14294  
QY 2146 CAGGAGACACCAAGGGGAAGTCCCAAGAGACCAGCAGCGTGGCATTTGGACAGGGA 2205  
Db 14293 NNN 14234  
QY 2206 GGTCTGTCTGGACAGGTGCTTTGGATAAGGGAGGAAATGGTGCACTTCCATCCTCCTC 2265  
Db 14233 NNN 14174  
QY 2266 CTTCTCTGTTCAACCTCTAACTACATGGGGCACAGGACCAGTGGGACTCCATAATGA 2325  
Db 14173 NNN 14114  
QY 2326 TGGGATGGGTGGATGGAAGGAAGGAGGAACAACACTTTCATTCCTGTTATTT 2385  
Db 14113 NNN 14054  
QY 2386 ACAGAACAGCCAGGTGCGGTGCTCAGCTTGCCATTTAGCACCTTTGGGAGGCTGAGGT 2445  
Db 14053 NNN 13994  
QY 2446 GGGTGAATTACCTCAGTCAGGAGTTCAAGACAGCCCTAGACACGCTAGAGAACC--- 2502  
Db 13993 GGGTGAATTTCTTGAGCCAGGGGTTCAAGACCAGCTTGGGCAATATGTTGGAACCCCTGC 13934  
QY 2503 ---CATCTCTACTAAGATATAAATAGCTGGCGTAGTGGCATATGCTCTGTAATCCCA 2559  
Db 13933 CTGACTCTATAAAAAAATACAAAATTAGCCAGGAGTGGTGTGTTGATCTCTGTAGACCCA 13874  
QY 2560 GCTAGTCGGGAAGCTGAGGCAGGAGAATCGTTTGAACCCGAGAGGCGAGAGGTTGCGGTGA 2619

Db 13873 GCTACTTAGGAGGCTGAGGAGGAGGATTTGTTGAGTCCGCGAGGTCAAGGCTGCAGTGA 13814  
QY 2620 GCTGAGATCGGCGCATTCAGCTCCAGCCTGGGTGACAAAGCAAGACCTCGTCTCAATAAT 2679  
Db 13813 ACAATGATCAAGCACTGCTCCAGCTTGAGCGACAGTGGAGCACTGCTCTCAATAAAAA 13754  
QY 2680 AATAATAATTACAAAAACAGAA 2700  
Db 13753 CCCAAAAACCAACAAACAA 13733  
RESULT 8  
ABN95044/c  
ID ABN95044 standard; DNA; 110096 BP.  
XX  
AC ABN95044;  
XX  
DT 13-AUG-2002 (first entry)  
XX  
DE Gene #1542 used to diagnose liver cancer.  
XX  
KW Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;  
metastatic liver tumour; cytostatic; expression profile; disease state;  
disease progression; drug toxicity; drug efficacy; drug metabolism.  
XX  
OS Homo sapiens.  
XX  
PN WO200229103-A2.  
XX  
PD 11-APR-2002.  
XX  
PF 02-OCT-2001; 2001WO-US030589.  
XX  
PR 02-OCT-2000; 2000US-0237054P.  
XX  
PA (GENE-) GENE LOGIC INC.  
XX  
PI Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;  
XX WPI; 2002-426119/45.  
XX  
PT Diagnosing and detecting the progression of liver cancer, hepatocellular  
carcinoma or metastatic liver tumor in a patient, involves detecting the  
level of expression of two or more genes in a liver tissue sample.  
XX  
PS Claim 1; SEQ ID NO 1542; 298pp; English.  
XX  
CC The invention relates to a novel method for diagnosing and detecting the  
progression of liver cancer, hepatocellular carcinoma or metastatic liver  
tumor in a patient, and differentiating metastatic liver cancer from  
hepatocellular carcinoma in a patient, involving detecting the level of  
expression of two or more genes represented in ABN93503-ABN97455 in a  
tissue sample. The method of the invention has hepatotropic, and  
cytostatic activity. The method is useful for diagnosing and detecting  
the progression of liver cancer, hepatocellular carcinoma and metastatic  
liver carcinoma in a patient. The method is useful for identifying  
expression profiles which serve as useful diagnostic markers as well as  
markers that can be used to monitor disease states, disease progression,  
drug toxicity, drug efficacy and drug metabolism. Note: The sequence data  
for this patent did not form part of the printed specification, but was  
obtained in electronic format directly from WIPO at  
ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 110096 BP; 25112 A; 29361 C; 28188 G; 27435 T; 0 U; 0 Other;

Query Match 8.3%; Score 247.6; DB 6; Length 110096;  
Best Local Similarity 47.3%; Pred. No. 8.7e-54;  
Matches 1149; Conservative 0; Mismatches 1224; Indels 57; Gaps 11;  
QY 287 CAAATGAGTGATAACAGTCCCATCAGCCGGGTGCGTAGTGGCTGAATCCCAA 346  
Db 58480 CAAACAAAAGAAAACAGTAAGGCTGCATGGGACGGTGCCTCACGCTGTAATCCAG 58421



QY 347 CACTTTGGAGCGGAGGTGGGTGATCACTTGAGTTCAGGATTCGAGACCGCCTGGC 406  
Db |||||  
58420 CAATTTGGAGGCCAGCGAGTGGATCACCTAGAGTCAGGATTCAGAGCAGCCTGGC 58361  
QY 407 CAACATGGTGAACCCCTGCTCTACTTAAATAATAAAATATAGTGGGCATGGTGGC 466  
Db |||||  
58360 CAACATGGTGAACCCCTATCTCTACTTAAATAATAAAATATAGTGGGCATGGTGGCAGG 58301  
QY 467 TACCTGTAAATCCGAGATCTTGGAGGTTGAGGCGAGGAAATCGCTTGAACCCGGGAGGC 526  
Db |||||  
58300 TGCCTGTAAATCCGAGTACTCAGGAGGCTGAGATGGGAAATCACTTGAACCTGGAGAT 58241  
QY 527 AGATGTTTGCAGTGAACCAAGACTGTGCCACTGCATCCAGTCTGGGCAACAGAGTGAGCC 586  
Db |||||  
58240 AGTGGTGCAGTGAGCCGAGATCTGTGGCAGTGCATCTCCCTGGGCGAGTGGAGAGAC 58181  
QY 587 TCCATCTCAAAACAAACAAACAAAGCAGTGGCCCATCATGTAGGATTCAGTGTAGTG 646  
Db |||||  
58180 CTCGTCTAAAAAAG 58121  
QY 647 AGGACTGAGCCTTGTGCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 706  
Db |||||  
58120 GCACCTGTAAATCCGAGTACTCAGGAGGCTGAGTGGGAGAAATTCCTTGAACCCAGGAGG 58061  
QY 707 ACCATCAATGATCCAGGTAAAGCCCTGAGGCTTCAGAAAGATGCCGAGGCTTTCAGG 766  
Db |||||  
58060 TGGAGGCTGCAGTGAGCCAGATCATGCCACTGCATCCAGCCTGGGCGACAGAGTGAGA 58001  
QY 767 TGTGGGGATTGGTGGCAAGCCCTCGAATAATAGAAACAGTCTCTGTATTACACAGA 826  
Db |||||  
58000 CCCCTTCTCAAAAAG 57941  
QY 827 AAGCAGAGCCCATGCTGGGTGCTGCCAGAACTCAGTAGTAATTAAGCAGCAGCCGGT 886  
Db |||||  
57940 GGG-TGGAGGCCAGGGCCCTGGAAAGGCCCTGTATGCTGTGTGAAGAAATTTAGACTTGT 57882  
QY 887 GCTGCTTCCCAGCGCACCTAGGCCAGTGGGAAACAG-----ACTCACCACAGT 938  
Db |||||  
57881 CTGTGCTGGTGGGATGGTTTGTAGCTTAACAACAGAGAGTAGGAATCAGATTACAGT 57822  
QY 939 CCCAGCCAGAGTGTTCAGGCCCAAGATGGGAGACGCGGGA---GAAAGGTCAAGGTG 995  
Db |||||  
57821 GTAGAATGTCACTAGGGGCTGACACTGTGGCTCAGCTTGTATCTCAGCACTTTGGGG 57762  
QY 996 GATGGGGAGGGTTCAGGCAAGAGGGGTGAGGCCAGGCTGAGGAGAGCCCTGGGACTG 1055  
Db |||||  
57761 CCAGGCAGGATTTTGTAGGCCAGGAGTTTGAACACAGCTTGGGTAAACATAGCAAGACC 57702  
QY 1056 TAGGAATTTAGAGAGGTACTGTACCCCGCATGTTTGTGAGGAGAGATTTCAGGAAGTCT 1115  
Db |||||  
57701 TCCATCTATACAGAAATTAACAATAATTAGCCAGAGTAGTGGTACCGCTGTAGT 57642  
QY 1116 CTGGAAGAGAGGCTGTCGAGCTGAGA--CTCATAGATGAGTGGGGAGGGGTGTTCCAG 1173  
Db |||||  
57641 CCTAGATATTTCAGGAGCAGACGAGAGAGATCACTTGAGCCTGGGGGATTCAGAGTTG 57582  
QY 1174 GCAGAAAGCCAGCCTACAAAGCATGACTTTGAGAAAGCATTCATCCATTCACATG 1233  
Db |||||  
57581 CAGTAATTGTACTGTACTCCAGCTTGGCAATGTAGCAAAAGCCCTGCCCTTCAAAAGAAA 57522  
QY 1234 ATGAATTTTCAGACTGGGACGCTGGCTCATGCTGTATCCAGACACTTTTGGAGGCTG 1293  
Db |||||  
57521 AAAAAGGTCAGCGGGTGGTGGCTCAGCTGTATTCAGCAGCACTTTGGAGGCGG 57462  
QY 1294 AATGGGAGGATGACTTAGCCTTAGGCATTTGTGACAAAGCCTGGGCAACATGGTGAAGCC 1353  
Db |||||  
57461 AGACGGCGGAT--CATGAGGTGAGCAGATTGAGACCATCTCTGGCTAACATGGTGAACC 57404  
QY 1354 CTCCTCCACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 1413  
Db |||||  
57403 CCGCCTCTACTAAAAATACAAAAAATTAGCCAGGTGTGGTGGCGGCGCTGTAGTCCCA 57344

QY 1414 GGTACATAGAAATGACTCAGCGAGATATGGTGT---CCTCTCTACTTGTGGAGAGGCGG 1470  
Db |||||  
57343 GCTACTCGGAGGCTGAGGCGAGGAGAAATGGATGAACCCGGGAGCGAGAGCTTGCAGTGA 57284  
QY 1471 GCTTATACTGCAGTAAGACAAATAGAGGAGGGAATATATCTCTAAATAGAGGTTACAGA 1530  
Db |||||  
57283 GCCGAGATTGTGCCACTGCATCCAGCCTGGGGACAGAGGAGACTCCATCTCAAAAAA 57224  
QY 1531 TTTGAGAGCAAAACACAGGGCACAGGCATATGTACAGGGTAAAGAGGGAATCAGGAAAG 1590  
Db |||||  
57223 AATATATATATATAAAGGTGAGCGTGGCTCATGCCATATATCCAGCACTTTGGG 57164  
QY 1591 CTTTCTCAGAGAGTGCATTTAAGCCGGGACATGAAGGATGAACGAGTTAGTTACCAA 1650  
Db |||||  
57163 AGGCCGAAGCAGGCGGA---TCACCTTGAGGTTCAGAGTTTGAGACCAGCCTGGGCAACAT 57107  
QY 1651 GGATGGGATGGAAGGGGTGAGGTGATGGAGGCGAGGGAACATGCAGGATCATAGGCT 1710  
Db |||||  
57106 GGTGAGACCTGTCTCTACTTAAATAATAAATAATAGCCGGCGTGTGTGTGTGCTG 57047  
QY 1711 AGACAGGGGATCTGACGCCCTTTCAGGAAGTGAAGAGAACACGCGAGTCTGTAGTGGT 1770  
Db |||||  
57046 TAGTCCCAAGTACTCAGAGGCTGGGCGACAGAAATGACTCGAGCCCAAGAGGTGGAGT 56987  
QY 1771 TAAGTAACAAAGCTGAGAACGCCAGGGAATCCCTGTGTATGCAGGGCTGTGAGTCACT 1830  
Db |||||  
56986 GGCAGTGAGCTGAGATTGCCACCTCCAGCCTTGAGCGACACTCCAGCCTGTCACTCCAGC 56927  
QY 1831 CAGAGTGTTTGGGCTTTTGTCTCTGGAGCAGTGCATTTTAAAGCAGGGAACAGCTGT 1890  
Db |||||  
56926 CTGAGCACCGAGTGAGAGTCTGTCTCAAAAAAATAAATAATAAATAAATAAATAAATAA 56876  
QY 1891 ATTACAGATTGGGAAGATCCCTGTGTGTCTGCCGAGGGAATGAACACTTGAGGCTAGGA 1950  
Db |||||  
56875 TTTTTCACAAAAGAAATTTCTTACAAAAGAAATAAATTTTAAACATAAATTTATT 56816  
QY 1951 GCCCAGGCTGATAGGAGGATCCAGGCTGATGGGAGGCTGGGAGGTCCGCGGTGATGGA 2010  
Db |||||  
56815 ACAAAATTTTTCACAAAATAAATAAAGTAATAAATAAATAAATAAATAAATAAATAAATAA 56756  
QY 2011 CCAGGCTGGGCGCAGGGGATGGGAGAGGAGTAAATTCGGAGAGGCTCGGGCTCTGG 2070  
Db |||||  
56755 GAAGTGAAGTTATAGAGTCTTAATGATGTAGATTTAGAGACATTAAGAGGTAGGGC 56696  
QY 2071 CCAGGAATGATGGTGGCTGAAAACAGGAGAGGAGAGATGCTTAGGCCACTTTGGAAC 2130  
Db |||||  
56695 CGGCGATGTGGCTCACGCTGTAATCCAGCACTTTGGGAGGCTGAGGCGAGGTAGA-- 56639  
QY 2131 ACAGTAGGCAAGGACAGGAGACACCCAGGGGAAGTCCCAAGAGACCAAGCAGCAGGCTG 2190  
Db |||||  
56638 TCAGAGGTGAGGAGTTGGAGACCAAGCCTGGCCAGATGTTGAACCCCTCTCTACTAA 56579  
QY 2191 GCATTGACAGGGAAGTCTGTCTGGAGCAGGTGTCTTGGATAAGGGAGGAATAATGTCG 2250  
Db |||||  
56578 AATACAAAATTTAGTGGTGTGGTGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 56519  
QY 2251 AGTTCCATCT 2310  
Db |||||  
56518 CGAGGCAAGAGATCG-----TTGAACCGGGAGTTGGAGGTTGAGGTGAGTGGTGGTGGTGG 56465  
QY 2311 GGAATCAATAATGATGGATGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2370  
Db |||||  
56464 CGCTCCACTGCACTCTAGCTTGGG-----CAACAAAGCAAGACTCTGT 56422  
QY 2371 TCATCTGTGTTATTTACAGAACAGGCGAGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2430  
Db |||||  
56421 CTCAAAAAATAAATAAATAAAGGCTGGGACAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 56362  
QY 2431 TTGGAGGCTGAGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2490  
Db |||||  
56361 TTGAGAGGCGGAGGAGGCGGATCACCTGAGGTGCGGAGTTTGGAGACCGCCTGAGCAAC 56302  
QY 2491 GTAGAGAAACCCCATCTCTACTGAAGATATAAATAATGAGTGGGCGTAGTGGCATATGCT 2550

Db	56301	ATGGAGAAATCCGCTCTCTACTATAAAATACAAAATCCCGAGTGTGGTACATGCCT	56242	PR	22-AUG-2000;	2000US-0226681P.
				PR	22-AUG-2000;	2000US-0226868P.
				PR	22-AUG-2000;	2000US-0227182P.
				PR	22-AUG-2000;	2000US-0227009P.
Qy	2551	GTAATCCAGCTAGTCGGGAAGCTGAGCGAGAGAAATCGCTTGAACCCGAGGAGGAGG	2610	PR	30-AUG-2000;	2000US-0228924P.
				PR	01-SEP-2000;	2000US-0229287P.
Db	56241	GTAATCCAGCTATTCGGGAGGCTGAGGCAAGAGAATCGCTTGAACCCGAGGAGGAGG	56182	PR	01-SEP-2000;	2000US-0229343P.
				PR	01-SEP-2000;	2000US-0229344P.
Qy	2611	TTGGGGTGAAGTGAATCGTGCCATTGCACTCCAGCCTGGGTGACAAAGCAAGACTCGT	2670	PR	01-SEP-2000;	2000US-0229345P.
				PR	05-SEP-2000;	2000US-0229509P.
Db	56181	TTGCAGTGAAGCGAGATCGCGCATTTGCATCCAGCCTGGGCAATGAAGTGAAGTCAAACTCCA	56122	PR	05-SEP-2000;	2000US-0229513P.
				PR	06-SEP-2000;	2000US-0230437P.
Qy	2671	CTCAATAATAATAATAATTACAAAACAGAA	2700	PR	06-SEP-2000;	2000US-0230438P.
				PR	08-SEP-2000;	2000US-0231242P.
Db	56121	TCTGAAAAAAAATAATTAATAATAAAAA	56092	PR	08-SEP-2000;	2000US-0231243P.
				PR	08-SEP-2000;	2000US-0231244P.
				PR	08-SEP-2000;	2000US-0231413P.
				PR	08-SEP-2000;	2000US-0231414P.
				PR	08-SEP-2000;	2000US-0232080P.
				PR	08-SEP-2000;	2000US-0232081P.
				PR	12-SEP-2000;	2000US-0231968P.
				PR	14-SEP-2000;	2000US-0232397P.
				PR	14-SEP-2000;	2000US-0232398P.
				PR	14-SEP-2000;	2000US-0232399P.
				PR	14-SEP-2000;	2000US-0232400P.
				PR	14-SEP-2000;	2000US-0232401P.
				PR	14-SEP-2000;	2000US-0233063P.
				PR	14-SEP-2000;	2000US-0233064P.
				PR	14-SEP-2000;	2000US-0233065P.
				PR	21-SEP-2000;	2000US-0234223P.
				PR	21-SEP-2000;	2000US-0234274P.
				PR	25-SEP-2000;	2000US-0234997P.
				PR	25-SEP-2000;	2000US-0234998P.
				PR	26-SEP-2000;	2000US-0235484P.
				PR	27-SEP-2000;	2000US-0235834P.
				PR	27-SEP-2000;	2000US-0235836P.
				PR	29-SEP-2000;	2000US-0236327P.
				PR	29-SEP-2000;	2000US-0236367P.
				PR	29-SEP-2000;	2000US-0236368P.
				PR	29-SEP-2000;	2000US-0236369P.
				PR	29-SEP-2000;	2000US-0236370P.
				PR	02-OCT-2000;	2000US-0236802P.
				PR	02-OCT-2000;	2000US-0237037P.
				PR	02-OCT-2000;	2000US-0237038P.
				PR	02-OCT-2000;	2000US-0237039P.
				PR	02-OCT-2000;	2000US-0237040P.
				PR	13-OCT-2000;	2000US-0239935P.
				PR	13-OCT-2000;	2000US-0239937P.
				PR	20-OCT-2000;	2000US-0240960P.
				PR	20-OCT-2000;	2000US-0241785P.
				PR	20-OCT-2000;	2000US-0241786P.
				PR	20-OCT-2000;	2000US-0241787P.
				PR	20-OCT-2000;	2000US-0241808P.
				PR	20-OCT-2000;	2000US-0241809P.
				PR	20-OCT-2000;	2000US-0241826P.
				PR	20-OCT-2000;	2000US-0242221P.
				PR	01-NOV-2000;	2000US-0244617P.
				PR	08-NOV-2000;	2000US-0246474P.
				PR	08-NOV-2000;	2000US-0246475P.
				PR	08-NOV-2000;	2000US-0246476P.
				PR	08-NOV-2000;	2000US-0246477P.
				PR	08-NOV-2000;	2000US-0246478P.
				PR	08-NOV-2000;	2000US-0246523P.
				PR	08-NOV-2000;	2000US-0246524P.
				PR	08-NOV-2000;	2000US-0246525P.
				PR	08-NOV-2000;	2000US-0246526P.
				PR	08-NOV-2000;	2000US-0246527P.
				PR	08-NOV-2000;	2000US-0246528P.
				PR	08-NOV-2000;	2000US-0246532P.
				PR	08-NOV-2000;	2000US-0246609P.
				PR	08-NOV-2000;	2000US-0246610P.
				PR	08-NOV-2000;	2000US-0246611P.
				PR	08-NOV-2000;	2000US-0246613P.





isolated from a range of human tissues disclosed in the specification.  
The nucleic acids, proteins, antibodies and (ant)agonists are useful in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and ovarian cancer and other cancers of the adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune disorders e.g. Addison's disease, allergies, autoimmune haemolytic anaemia, autoimmune rheumatoid arthritis, diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c) cardiovascular disorders such as myocardial ischaemia; (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f) infectious diseases such as viral, bacterial, fungal and parasitic infections. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

Sequence 5041 BP; 1480 A; 1074 C; 1103 G; 1384 T; 0 U; 0 Other;

Query Match 8.2%; Score 247.2; DB 5; Length 5041;  
Best Local Similarity 65.9%; Pred. No. 3.2e-54;  
Matches 407; Conservative 0; Mismatches 203; Indels 8; Gaps 3;

QY 2392 CAGCCAGGTGCGGTGCTCAGCTTGGCATTCTAGCATTGGGAGGTGAGGTGGTGG 2451  
DB 3004 CAGCCGGGCGCAGTGGCTCATGCTTAATCCAGCACTTTGGGAGGCGGAGCGGAG 3063  
QY 2452 ATTACCTCAGTCAAGGAGTTCAAGACAGCCTAGACACGTAGAGAACCCCATCTCTAC 2511  
DB 3064 ATGACTTGAGTCAAGAGTTCGACACAGCCTGGCCACGTGGCAACCCCATCTCTAC 3123  
QY 2512 TGAAGAT-ATAAATTAGCTGGGCTAGTGGCATATGCTGTATCCAGCTAGTCGGGA 2570  
DB 3124 TAAATAAAGAAATTAGCCGGGTGATGSCACATGCTGTATCCAGCTACTCAGGA 3183  
QY 2571 AGCTGAGCAGGAGAAATCGCTTGAAACCGAGAGCAGAGTTGGGTGAGCTGAGATCGT 2630  
DB 3184 GGTGAGCAGGAGAAATCGCTTGAGCCCGCAGGAGCAGAGATTGAGTGAAGATCGC 3243  
QY 2631 GCATTCGACTCCAGCTGGGTGACAAAGCAAGACCTCGTCTCAATATATATATATTA 2690  
DB 3244 ACCACTGCATCCAGCTCGGTGACAGCAAGACTCCATCTCAGAAAAAATAAAAAA 3303  
QY 2691 CAAATC-----AGAAGGAGCCTGGGTCTATCCAGCTACTCTTTTCAGGAGATGTACT 2746  
DB 3304 AAAAACTTTCTCTCGTCCCATCTTCTTCGCTCTCTGACTTCTTCTGACTGCTGTT 3363  
QY 2747 CTTTACCAAGGCAAGAGTGGAGAACAGTTTGAATATGCAATTTATTGAGCACTAC 2806  
DB 3364 GGAAGTCTTAATATGAAAGAGGCTGAAACTAGGCATTAATAATTAATCTGATGCTCAAAGT 3423  
QY 2807 TGAGTCTCATCTCCCTGGGCTAGGCTGGAATGGACTCAGATGGAGCCTG---AAGAGTCCC 2863  
DB 3424 AAGATTCCATCCAAACAGATCTCTTCTCCCTCCAGGGTCTTCCAGGAGTCTGCTA 3483  
QY 2864 CTTGAGGAACCTCACTAGAAAGAGGAGGAATGGCCGGGCGGTGGTCTCAGCCTGT 2923  
DB 3484 CTTATCCAAATAAAATGGTAAAGATAAAGACAGCCGGGCGGTGGTCTCAGCCTGT 3543  
QY 2924 AATCCCAACTTTGGAGCCTGAGTGGGTGATCAAGGTCAGGAGATCGAGACCAT 2983  
DB 3544 AATCCAGCACTTTGGAGGCGGAGGCGGAGGATCAAGAGTCAAGAGATCGAGACCAT 3603  
QY 2984 CTTGGCTTAACACAGTGAA 3001  
DB 3604 CTTGGCTTAACAGTGAA 3621

RESULT 11  
ABQ74964\_5/c  
Continuation (6 of 8) of ABQ74964 from base 500001 (Human kinase protein genomic DNA seq  
WP Sequence split into 8 fragments LOCUS ABQ74964 Accession ABQ74964  
WP Fragment Name Begin End  
WP ABQ74964\_0 1 110000  
WP ABQ74964\_1 100001 210000

WP ABQ74964\_2 200001 310000  
WP ABQ74964\_3 300001 410000  
WP ABQ74964\_4 400001 510000  
WP ABQ74964\_5 500001 610000  
WP ABQ74964\_6 600001 710000  
WP ABQ74964\_7 700001 786431

Query Match 8.2%; Score 245.8; DB 6; Length 110000;  
Best Local Similarity 84.2%; Pred. No. 2.6e-53;  
Matches 277; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY 282 TCTGTCAATGAGCTGATAACAGTGGCCATCAGCGGGTGCAGTGGCTAGTGCCTGAAT 341  
DB 72672 TTTTTCATTACTGATTAACTGACTAGAGGCCAGTGTGTGGCTCATGCTTAAAT 72613  
QY 342 CCCAACACTTTGGGAGGCGGAGTGGTGGATCACTTGGAGTCCAGAGTTCGAGACCAGC 401  
DB 72612 CCCAGCACTTTGGGAGGCGGAGGTAGGTGGATCACTTGGAGTTCGAGACCAGC 72553  
QY 402 CTGCCCAACATGTGAAACCCCTGTCTCTAATAAAAAATATAAAAAATTAGCTGGCATGGT 461  
DB 72552 CTGCCCAACATGTGAAACCCCTGTCTCTAATAAAAAATACAAAAATTAGCTGGATGGT 72493  
QY 462 GTGCGTACCTGTATCCAGATCTTCGGAGGTTTGGCAGGAGATCGCTTGAACCCGG 521  
DB 72492 CGGTACACCTGTGTATCCAGTCTTGGGAGGTTTGGGAGGATTAATCACTTGAACCTGG 72433  
QY 522 GAGCAGATGTTCAGTGAACCAAGACTGTGCCACTGCACCTCCAGTCTGGGCAACAGAGT 581  
DB 72432 GAGGTGAGGTTGCACTGAGTCAAGATCGTGGCACTGCACCTCCAGCTGGGTGACAGT 72373  
QY 582 GAGCTTCATCTCAACAAACAAACAAAA 610  
DB 72372 GAAACTCCGTCTCAACAAACAAACAAAA 72344

RESULT 12  
ADC85367/c  
ID ADC85367 standard; DNA; 96593 BP.  
XX  
AC ADC85367;  
XX  
DT 01-JAN-2004 (first entry)  
XX  
DE Mouse Bln coding sequence.  
XX  
KW Cytostatic; gene therapy; vaccine; cancer; carcinoma-associated gene; CA;  
KW secreted; transmembrane; intracellular; ds.  
XX  
OS Mus sp.  
XX  
PN WO2003045230-A2.  
XX  
PD 05-JUN-2003.  
XX  
PP 02-DEC-2002; 2002WO-US038582.  
XX  
PR 30-NOV-2001; 2001US-00997722.  
XX  
PA (SAGR-) SAGRES DISCOVERY.  
XX  
PI Morris DW, Engelhard EK;  
XX  
DR WPI; 2003-513603/48.  
XX  
PT New recombinant nucleic acid comprising a nucleotide sequence of any of  
PT the carcinoma-associated (CA) genes, useful for screening for drug  
PT candidates for diagnosing or treating carcinomas.  
XX  
PS Claim 1; SEQ ID NO 153; 983pp; English.  
XX  
CC The invention relates to a recombinant nucleic acid comprising a  
CC nucleotide sequence selected from any of the fully defined carcinoma-

CC associated (CA) genes from the 50 tables given in the specification. The  
 CC CA proteins are secreted, transmembrane or intracellular proteins. The  
 CC recombinant nucleic acids are useful for screening for drug candidates  
 CC for diagnosing or treating carcinomas. Sequences given in ADC85215-  
 CC ADC85514 represent CA genes of the invention.  
 XX  
 SQ Sequence 96593 BP; 27724 A; 19526 C; 19631 G; 29712 T; 0 U; 0 Other;  
 Query Match 8.2%; Score 245.6; DB 9; Length 96593;  
 Best Local Similarity 48.2%; Pred. No. 2.8e-53;  
 Matches 1158; Conservative 0; Mismatches 1174; Indels 68; Gaps 14;  
 QY 314 GCCGGTGCAGTGGCTAGTGCCTGAAATCCCAACACTTTGGAGGGGAGGTGGTGGAT 373  
 DB 17837 GCCGGCGTGGTGGCTTATGCTCTAATCCCGACCTTTGGAGGCGGAGACGGGAGAT 17778  
 QY 374 CACTTGAGGTGAGGAGTTCGAGACACAGCTGCGCAACATGCTGAAACCCCTGCTCTACTA 433  
 DB 17777 CACTTGAGGTGAGGAGTTCGAGACCATCTGCAACACATGGCCAAACCGCTCTCTACTA 17718  
 QY 434 AAAATATAAAATTAGCTGGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 493  
 DB 17717 AAAATACAAAATTAGCCAGGATGATGCTGACACCTATAACCCAGCTATTTCAGTAGG 17658  
 QY 494 TTGAGCAGGAGAAATCGCTTGAAACCGGAGGAGCAGATGTTGAGTGAACCAAGACTGTGC 553  
 DB 17657 CTGAGACATGAGAAATCTCTTGAACCCAGCTGTAGAGGTTCAGTGAGCCNAGATCATAC 17598  
 QY 554 CACTGCACCTCCAGTCTGGGCAACAGAGTGAAGCTCCATCTCAAAACAAACAAACAAAGC 613  
 DB 17597 CACTGCACCTCCAGCTGGGCAACAGAGTGAAGCTCCCTCTCAAAACAAACAAACAAA --- 17541  
 QY 614 AGTGCCCATCATGTAGATTTGAGTGTGATGAGTGAAGTGAAGCTTGTGCAAAAGTGAGCA 673  
 DB 17540 -----CTTCCTCTGGGTAGGACAAAGGGGCTGGGAGTGAATTAATCAACCAATGGCCAA 17487  
 QY 674 CTCACATAATC-ACCAGGTTGATGATCAGTGAATCAATCAATGATCCAGGTAAAGCCCT 732  
 DB 17486 TTATTTAATCAATCATGTCTAGTAAAGAGGCTCAGTAAACCCAGAGGATAGGCT 17427  
 QY 733 GAGGGTTCAGAAAGATGCGGAGCGCTTTCAAGGTGCTGGGATTTGGTGGCAAGCCCTC 792  
 DB 17426 CAGGAGGCTCTAGGTTGGTGAGAAATGTAGAGATGTTGGAGGCTGTGTGCTGGAG 17367  
 QY 793 GAATAATAGAAACAGTTCTCTGTATTAACAAGAAAGAGAGGCCATGCTGGGTGCTG 852  
 DB 17366 AGGGCATGGAAGTGCCACACCCCTCCACATACCTTGTCTGTGTATTTCTCCAGCTG 17307  
 QY 853 CCAGGAACCTCAGTAGTAATCTAGACAGCACCGGTGCTCTCCACAGCGCACCTAGGCCA 912  
 DB 17306 GCTGCTTCTGAGTTGTTATTAATAGCTGATATCTAGTAAGGAATGTTCTGTGACTG 17247  
 QY 913 GTGGGAAACAGACTCACACACAGTCCAGGCCAGAGTGGTTCAGGGCCCAAGATGGGAA 972  
 DB 17246 AATGTGGAGACTTACCACCTTGCTGTGTAGTCTGAAAGTGGGGGTCTGGAATGGGGTG 17187  
 QY 973 GCACGGGGAAGAGGTCAAGGTGGATGGGAGGGGTGAGGCAAGAGGGGTGAGGGCCA 1032  
 DB 17186 GACAGGGAGGACTTAACATTTGGGGTCTGCACCTAACACCTGGCAGTTAGTGTGACAGCCTG 17127  
 QY 1033 GGTGAGGGAAGCCCTGGGACTGTAGGAATTTAGAGGAGGTACTCTACCGCGCATGTTG 1092  
 DB 17126 AGCTAAATGTAGACACCCAGTGTCTGCTAAGAACTGAACTATTTGGGGAAA----- 17073  
 QY 1093 GTGAGGAGATTCAGGAAGTCTTCTGGAAGAGAGCTGTCGAGGCTGAGACTCATAGA 1152  
 DB 17072 ---AAAAACACCACACATTTGGTGACAGAAGTGAAGAAATATTATATGGGTACAGAGTAG 17016  
 QY 1153 TGAGTGGGAGGTTGTTCCAGGAGCAAGACAGCACCTACAAAGCATGACTTTGAGAG 1212  
 DB 17015 AGAGGAAATGGAGTTTCTTTACAGTGGTATACGCA-----TAGAATG 16973  
 QY 1213 AAGCAATTCATCCATTCACTGATGAATTTTTCAGACTGGGACGCTGGCTCATGCTGTAA 1272

DB 16972 AAATATTGAGCTTTAAAGAAATAATTT--GACTAGATGAGTGGCTCATGCTGTATA 16915  
 QY 1273 TCCAGCACTTTGGAAAGCTGAATGGGAGGATGATCTTGGAGCTAGGCAATTTGTGACAA 1332  
 DB 16914 TCCAGCACTTTGGAGGCAAGGAGGAGTATCGCTTGGAGCCAGGAGTTTGTGAGACCA 16855  
 QY 1333 CCTGGGCAACATGTTGAGACCC--TGCCCTCCACAAACAAACAAACAAACAAATCAT 1390  
 DB 16854 CTTAGGCAACAAAGCAAGACCCAGTATCTTACAAATAATAAAAGTTAGCCAGCATGTT 16795  
 QY 1391 TATACCTGCTACATGAGGTACCAAGTATCATAGAAATGACTCAGGCAGATATGTTGCTCT 1450  
 DB 16794 GGTACACT 16735  
 QY 1451 TCTTACTGTGGAGAGGCGGCTTATCTGCTAGTGAAGCAATAGAGGAGGGAATATAT 1510  
 DB 16734 GGAGTTAGAGGCTGAGTGAAGCAAGCATCTTCTTCTCTCTCTCTCTCTCTCTCTCTCT 16675  
 QY 1511 CCTAAATGAGAGGTACAGATTTGAGAGCAACACAGGGCAAGGCATATGTACGAGGCT 1570  
 DB 16674 CAAGACCCCTGCTGAAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 16615  
 QY 1571 AAAGAGGGAATCAGGAGAGGCTTCTCAGAGAAGTGACATTTTAAGCC-----GGGACATG 1625  
 DB 16614 AAAAATAATCTCGCAATGAGATGACACAGATGAACCTTGAAGCCATTAAAGCAAGTG 16555  
 QY 1626 AAGGATGAACAGATTAGTTTCCACAGGATGGATGAAGGGGTGAGATGATGAGGCA 1685  
 DB 16554 AAGTAAGCCAGTCACAGGACAAATACTATGATTTCTCACTGCAACAGGAGTATCTAAATA 16495  
 QY 1686 GAGGGAATGCAAGGATCATAGGCTTAGACAGGGGATCTTGACGCCCTTCAGGAAGTGAGA 1745  
 DB 16494 GGCAACTCATAGAACAAAGAGGGGCAAGTGTGCAAGGTAGGGCTTTCAGGAAA 16435  
 QY 1746 GAAGACCCAGCGAGTGTAGTGGGTTAAGTAACAAAGCTGAGAAGCCAGGGAATCCCTG 1805  
 DB 16434 TGGGAGGTTGTTAATCAAAAGATATAAAGTTTCAGTAATGCAAAATGAACGCTTCTAGA 16375  
 QY 1806 GTATGAGGCGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1865  
 DB 16374 GATCTGCCATACAACATTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 16316  
 QY 1866 TCGATTTTAAAGCAGGCAACAGCTGATTCAGAGTTGGGAGATCTCTGCTGCTGCTGCTG 1925  
 DB 16315 TTAAGAAAACAAATTTTAAATAATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT 16256  
 QY 1926 AAGGGGATGAACTGAGGCTAGGAGCCAGGGGTGATGGGAGGATCCAGGGGTGATGGGG 1985  
 DB 16255 GCTAGAGCGAGTGGGACAAATCATAGTTCACTGCAAGCTCAGACTCTTAAGCTCAAGTGA 16196  
 QY 1986 AGCTGGGAGGCTCGCGGTGATGGACCAAG--GCTGGGGCCAGGGGATGGGAGGAGG 2042  
 DB 16195 TCTCTCTGCTCAGGCTTGGAGAACTCTGGAATCAGGTGATGATGATGATGATGATGATG 16136  
 QY 2043 AGTAATTGGGAGGCGCTCGGGCTCTGGCCGAGGAATG-----GATGTGGGCTGAAAC 2096  
 DB 16135 AATTTTAAATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 16076  
 QY 2097 AGGAGGAGGAGATGCTTTAGGCCACTTTTGGAAACACAGTAGGGCAAGGACAGGACACC 2156  
 DB 16075 CTGGCTCAAGAGATCTTGTCTCTCAGGCTCTCCAAAGTGTGGGATTCAGAAAGTGAGC 16016  
 QY 2157 CAAGGGAGGAGTCCCAAGAGACCAAGCAGAGGCTGGCATTTGGACAGGAGGCTGCTGCTG 2216  
 DB 16015 TACCATGCCAGCAAAATTTACTTTTAAATAATAATAATAATAATAATAATAATAATAATA 15956  
 QY 2217 AGCAGGTGCTTGGATAAGGGAGGAAATGGTGC-----AGTTCATCTCTCTCTCTCTCT 2267  
 DB 15955 ACAACGTGGAGAAATCTCAAAACATAATAACGCAAACTAGAAAGACTCAGTAGAATGCA 15896  
 QY 2268 TCTCTCTCAACCTCTTAAACTATCATGGGGCAAGGCCCGAGCTGGGACTCCATTAATCATG 2327



Db 15895 TATAGTATTATCCCAATTTACAGAGCTCAAAACAGGCAAACTGAGCTATATATTTAT- 15837  
QY 2328 GGATGGGTGGATGGAGGAAGGAGGAAACAACACTCTTCAATTCCTGTTATTTAC 2387  
Db 15836 AGATGTATAGATCAGGGCAAACTATGAAGACAAAGAAAAGCAATTATCACAATAATTGGGAC 15777  
QY 2388 AGAACAGGCCAGGTGCGGTGCTCAGCTTGCCATTCTAGCACTTTGGGAGGCTGAGGTGG 2447  
Db 15776 AATAGGCCAGAGCGGTGGCTCATGCTGTAAATCCAGCACTTTGGGAGGCCGAGGAG 15717  
QY 2448 GTGATTAACCTCAGGTCAAGGTTCACAGCCAGCTGACAAACGTVAGAGAAACCCCATCT 2507  
Db 15716 GCATTAACCTCAGGTCAAGGTTCACAGCCAGCTGACAAACGTVAGAGAAACCCCATCT 15657  
QY 2508 CTACTGAGAT--ATATAATTAGCTGGCGGTAGTGGCATATGCCCTGTAAATCCAGCACTAGT 2565  
Db 15656 CTACTAAAAATAGAAAAAATTAGCCAGCATGGTGGCAGGCGCTGTAAATCCAGCACTACT 15597  
QY 2566 CGGGAAGCTCAGGCAGGAGAAATCGCTTGAACCCGAGAGGCGAGAGTTCCGGTGTAGCTGAG 2625  
Db 15596 TGGGAGGCTCAGGCAGGAGAAATGCTTGAACCCAGGAGAGAGAGTTGCACTGAGCCGAG 15537  
QY 2626 ATCGTGCAATTGCACTCAGCCTCGGTGACAA-AGCAAGACCTCGTCTCAATAATAATAA 2684  
Db 15536 ATCGCGCACTGCACCTCAGAAATGGCAACCAAGAGCAAACTCCGCTCTCAAAAATAATAA 15477

## RESULT 13

ADA02888/c  
ID ADA02888 standard; DNA; 96594 BP.  
XX ADA02888;  
XX  
XX  
XX 06-NOV-2003 (first entry)  
XX Human BLM carcinoma associated gene, SEQ ID NO:1406.  
XX Human; carcinoma associated; oncogene; carcinoma; cancer; breast;  
XX prostate; lymphoma; leukaemia; cytostatic; gene therapy; drug screening;  
XX Gene; ds.  
XX Homo sapiens.  
XX OS  
XX WO2003057146-A2.  
XX PN  
XX PD 17-JUL-2003.  
XX  
XX 26-DEC-2002; 2002WO-US041414.  
XX 26-DEC-2001; 2001US-00035832.  
XX (SAGR-) SAGRES DISCOVERY.  
XX PA  
XX Morris DW;  
XX PI  
XX WPI; 2003-587068/55.  
XX DR  
XX  
XX PT New recombinant nucleic acid encoding carcinoma associated protein,  
XX useful for preparing compositions for treating carcinomas.  
XX  
XX PS  
XX PS Claim 1; SEQ ID NO 1406; 245pp; English.  
XX  
XX The invention relates to recombinant carcinoma associated (CA) nucleic  
XX acid sequences from mouse and human (ADA01482-ADA03094), and to  
XX recombinant carcinoma associated proteins (CAP) encoded by them. The  
XX invention also encompasses expression vectors and host cells comprising a  
XX CA nucleic acid, a polypeptide (especially an antibody) that specifically  
XX binds to the protein, and a biochip comprising CA nucleic acid or  
XX fragments thereof. The sequences of the invention were identified using  
XX oncogenic retroviruses, which insert into the genome of the host organism  
XX at random. Many of these do not carry transduced host oncogenes or  
XX pathogenic trans-acting viral genes, meaning that cancer incidence is a  
XX direct consequence of the effects of proviral integration into host

CC protooncogenes. The CA nucleic acid sequences can be used to diagnose  
CC carcinoma (especially breast cancer, prostate cancer, lymphoma or  
CC leukaemia) or a propensity to carcinoma by determination of the sequence  
CC of a CA gene, or by determination of CA gene expression in particular  
CC tissues. CA nucleic acids, proteins and antibodies are also useful as  
CC therapeutic agents and in screening and evaluating drug candidates. The  
CC present sequence represents a specifically claimed human CA nucleic acid  
CC sequence of the invention. Note: The complete sequence data for this  
CC patent did not form part of the printed specification, but was obtained  
CC in electronic format directly from WIPO at  
XX ftp.wipo.int/pub/published\_pct\_sequences.

Seq sequence 96594 BP; 27725 A; 19526 C; 19631 G; 29712 T; 0 U; 0 Other;

Query Match 8.2%; Score 245.6; DB 8; Length 96594;  
Best Local Similarity 48.2%; Pred. No. 2.8e-53;  
Matches 1158; Conservative 0; Mismatches 1174; Indels 68; Gaps 14;

QY 314 GCCGGGTGCAGTGGCTAGTGCCTGAATCCCAACACTTTGGGAGGCGGAGTGGGTGGAT 373  
Db 17837 GCCGGGCGTGGTGGCTTATGCCCTGTAATCCAGCACTTTGGGAGGCGGAGCGGAGAT 17778  
QY 374 CACTTTGAGGTCAAGAGTTTCGAGACAGCCTGGCCAAACATGCTGAAACCCCTGTCTCTACTA 433  
Db 17777 CACCTGAGGTCAAGAGTTTGAGACCATCTCGACACATGTCGACAAACGCTCTCTCTACTA 17718  
QY 434 AAAATATAAAATTTAGTGGGCATGGTGGTCGTAACCTGTATATCCAGATATCTTGGGAGG 493  
Db 17717 AAAATACAAAAATTAGCCAGCATGATGGTGCACACCTATTAACCCAGCTATTTCAGTAGG 17658  
QY 494 TTGAGGCGAGGAGATCGCTTGAACCCGGAGGAGAGATGTTTCAGTGAACCAAGACTGTGC 553  
Db 17657 CTGAGACATGAGAAATCTCTTGAACCCAGCTGGTAGAGTTTGCACTGAGCCCAAGATCATAC 17598  
QY 554 CACTGCACTCCAGTCTGGGCAACAGAGTAGAGCTCCATCTCAAAACAAACAAACAAAGC 613  
Db 17597 CACTGCACTCCAGCTGGGCAACAGAGTAGAGCTCTCTCTCAAAACAAACAAACAAACAA 17541  
QY 614 AGTCCCATCATGTAGGATTTAGTGTAGTGTAGGAGCTGAGCCCTTTGTGCAAAAGTAGCA 673  
Db 17540 -----CTTCTCTGGGTAGGACAAAGGGCTGGGGAGTGACTTAATCAACCAATGGCCAA 17487  
QY 674 CTCACCTAATC-ACCAGGTTGTAGTATCAGTGTATTAACCATCAATGATCCAGGTAAGCCCT 732  
Db 17486 TTATTTAATCAATCATGTCTAGTAGGAGGCCCTACGTAAAAACCCAGAGGATAGGGCT 17427  
QY 733 GAGGTTTCAAGAAAGATGCCGAGCGCTTTCAAGGTGCTGGGGATTGTTGGGCAAGCCCTC 792  
Db 17426 CAGGAGCTTCTAGTTGGTGAGAACATGTAGAGATGGTGGAGGCTGGTGTGCTGGAG 17367  
QY 793 GAATAATAGAAACAGTTTCTCTGTATTACAAAGAGCAGAGGCCCATGTCTGGGTGCTG 852  
Db 17366 AGGCGATGGAAGTGCCACACCCCTTCCACATACCTTTGCTGTGTATTCTTCCAGCTG 17307  
QY 853 CCAGGAACCTCAGTAGTAACCTAAGACAGCACCGGTGCTGCTTCCCGGCGACCTAGGCCA 912  
Db 17306 GCTGCTTCTCAGTTGTTTATAATAAGCTGAATACTTAGTAAGGAAACTGTTCTGTGACTG 17247  
QY 913 GTGGGAAACAGACTCACACACAGTCCAGCCAGAGTGGTCAGGCGCCCAAGATGGGAA 972  
Db 17246 AATGTGGAGACTTACCACCTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 17187  
QY 973 GCACGGGGAGAAAGGTCAAGGTGGGATGGGAGGGGTTCAGGGCAAGAGGGGTTCAGGGCCA 1032  
Db 17186 GACAGGGAGACTTAACATTTGGGGTCTGCACCTAAGCACTGGCAGTTAGTGTTCAGACCTG 17127  
QY 1033 GCGTGAAGGAGCCCTGGGACTGTAGAAATTTAGAGAGAGTACTGACACCGGCAATGTTG 1092  
Db 17126 AGCTAAATGTAGACACCCAGTGTCTGCTAAAGAACTGAACCTATTGTGGGGAAA----- 17073  
QY 1093 GTGAGGAGATTTCAGGAAGTCTTCTCTGGAGAGAGAGCTGTGCGGAGCTGAGACTCATAGA 1152  
Db 17072 ----AAAACACACACATTTGGTGACAGAAGTGAGGAATTTATTATATGGGTGACAGAGTAG 17016

QY 1153 TGAATGGGAGGGTGTTCAGGCGAAGAACAGACGACCTACAAAGCATGACTTTGAGAG 1212  
 Db |||||  
 QY 17015 AGAGGAAATGGAGTGTTCCTTTACAGCTGGTATACGCA-----TAGAATG 16973  
 Db |||||  
 QY 1213 AAGCATTCATCCATCACTGATGAATTTTCAGACTGGGCGACGCTGGCTCATGCTGTAA 1272  
 Db |||||  
 QY 16972 AAATATTGAGCCTTTTAAAGAAATAATTT--GACTAGATGAGTGGCTCATGCTGTAA 16915  
 Db |||||  
 QY 1273 TCCAGCAGCTTTGGAAGGCTGAATGGGAGGATGACTTTGAGCCTAGGCCATTTTGTACAAG 1332  
 Db |||||  
 QY 16914 TCCAGCAGCTTTGGGAGGCCAAGGAGGAGTATCGCTTGGCCAGGAGTTTGGAGACCAG 16855  
 Db |||||  
 QY 1333 CTTGGGCAACATGTTGAGACCC--TGCCTCCACAAAACAAACAAACAAACAAACAAATCAT 1390  
 Db |||||  
 QY 16854 CCTAGGCAACAAAGCAAGACCCAGTATCTACAAAATAAATAAGTTAGCCAAAGCATGT 16795  
 Db |||||  
 QY 1391 TATACCTGTGTACATGGGTACCAGGTACATAGAAATGACTCAGGCGAGATATGGTGTCCCTC 1450  
 Db |||||  
 QY 16794 GGTCACTCCTGTGTCTCCTTCTCGGAGGCTGAGCGAGAGGATCATCTGAGCCCA 16735  
 Db |||||  
 QY 1451 TCCTACTGTGGAGAGCGGGCTTATACTGCACTAAGA CAATAGAGGGAGGGAATATAAT 1510  
 Db |||||  
 QY 16734 GGAGTTAGAGGCTGCAGTGAGCAACGATCATTTCTACTGCACTCCAGCCTGGATAGTGAG 16675  
 Db |||||  
 QY 1511 CCTAAAAATGAGAGGTACAGATTGAGAGCAAAACACAGGCGCACAGGCATATGTACGAGGCT 1570  
 Db |||||  
 QY 16674 CAAGACCTGTCTGAAAAATAATAAAATACAAATTAAGTTAAATAATAAAATAA 16615  
 Db |||||  
 QY 1571 AAAGAGGGAATCAGGGAAGGCTTCTCAGAAAGGTGACATTTAAAGCC-----GGGACATG 1625  
 Db |||||  
 QY 16614 AAAAAAATCTGCAACATGAGATGACACAGATGAACCTTGAAGCCATTAAAGCAAGTG 16555  
 Db |||||  
 QY 1626 AAGGATGAACGATTTAGTTTACCAAGATGGGATGGGATGGGATGAGATGATGGAGGCA 1685  
 Db |||||  
 QY 16554 AAGTAAGCCAGTACACAGGCAAAATACATGATATTCACCTGCAACGAGGTATCTAAATA 16495  
 Db |||||  
 QY 1686 GAGGGAACCTGCAGGATCATAGGCTAGACAGGGATCCTGAGCCCTTGGAGGAATGAGA 1745  
 Db |||||  
 QY 16494 GCAAACTCATGAAGCAAGAGGGGCACAGTGTGTCAGGGTAGGGGCTTTCAGGAAA 16435  
 Db |||||  
 QY 1746 GAAGACCGAGCTAGTGTGGTTAAGTAAACAAAGCTGAGAGCCAGGGAATCCCTG 1805  
 Db |||||  
 QY 16434 TGGGAGTGTATTAATCAAAAAGTATAAGTTTCAGTAATGCAAAATGAACGATTTCTAGA 16375  
 Db |||||  
 QY 1806 GTCATGAGGCGCTGTGATGATCAGTGTGAGTGTGGGCTTTGTTTCTCTGGGAGCAG 1865  
 Db |||||  
 QY 16374 GATCTGCCATACAACTTTGTGGC-TGTGTGGTACTATACATGTAAAACTCTGTTAAGAGG 16316  
 Db |||||  
 QY 1866 TCGATTTTAAGCAGGGAACAGCTGTATTCAGAGTTCGGGAAGATCCTGTGTTGCTGCCTG 1925  
 Db |||||  
 QY 16315 TTAAGAAAACAATTTTAAAAAATTTTTTTAAAGACAGAGTCTTGTCTGTCACTCAG 16256  
 Db |||||  
 QY 1926 AAGGGATGAACTGGAGGCTAGGAGCCAGGCTGATAGGAGGATCCAGGGTGTATGGG 1985  
 Db |||||  
 QY 16255 GCTAGAGCGAGTGGCACAATCATAGTTCACTGAGCCTCAGACTCCTTAAGCTCAAGTGA 16196  
 Db |||||  
 QY 1986 AGGCTGGAGGCTCGGCTGTATGGAACAGG--GCTGGGCGCAGGGGATGGGAGGAAG 2042  
 Db |||||  
 QY 16195 TCCTCTGCTCAGCCTCTGAGAACTGTGAACTACAGGTGTCATCATGCTCAACT 16136  
 Db |||||  
 QY 2043 AGTAATTTGGAGAGGCTGGGGCTCTGGCCGAGGAATG-----GATGGTGGGCTGAAAC 2096  
 Db |||||  
 QY 16135 AATTTTAAATTTTATAGACAAAGTCTCCATTTGTGCCAGACTGCTTGTGAATC 16076  
 Db |||||  
 QY 2097 AGGAGAGGAGATGCTTAGGCCACTTTGGAAACAGATAGGGCAAGGACAGAGACACC 2156  
 Db |||||  
 QY 16075 CTGGCTCAAGATCTTGTCTCTCAGCCTCCAAAAGTCTGGGATTAACAAGTGAGC 16016  
 Db |||||  
 QY 2157 CAAGGGAGTGCCCAAGACACAGCAGGCTGGCATTTGGACAGGAGGTTGTCTGG 2216  
 Db |||||  
 QY 16015 TACCATGCCCGCAAAATTTACTATTAAAAAATAAAGAAAGAAATAATAATCT 15956  
 Db |||||

QY 2217 AGCAGTGTCTTGGATAAGGGAGGAAAAATGGTGC-----AGTTCATCTCTCTCC 2267  
 Db |||||  
 QY 15955 ACAACGTGGAAGAACTCTACAAACATAATAACGCAAACTAGAAAGACTCAGTAGAATGCA 15896  
 Db |||||  
 QY 2268 TCTCTGTTCAACCTCTAAACTACATGGGGCACAGGACCCAGTGGGACTCCATAAATGATG 2327  
 Db |||||  
 QY 15895 TATAGTATTATCCCATTTACAGAGCTCAAAACCCAGGCAAACTGAGCTATATTATTAT- 15837  
 Db |||||  
 QY 2328 GGATGGGTGATGGAAGGAAGGAAGGAACAACTCTTCTATTCTCTGTTATTATAC 2387  
 Db |||||  
 QY 15836 AGATGTATAGATPCAGGGCAAACTATGAAGACAAGAAAAAGCAATTATCAGAAAAATGGGAC 15777  
 Db |||||  
 QY 2388 AGAACAGGCGCAGGTGCGGTGCTCAGCTTGCCTTCTAGCACTTTGGGAGGCTGAGGTGG 2447  
 Db |||||  
 QY 15776 AATAGGCCAGAGCGGTGGCTCATGCTGTAAATCCAGCACTTTGGGAGGCCGAGCGAG 15717  
 Db |||||  
 QY 2448 GTGGATTACCTCAGGTTCAGGAGTTCAAGACCAAGCCTTAGACAAACGTTAGAGAAAACCCCATCT 2507  
 Db |||||  
 QY 15716 GCAGATTACCTGAGGTTCAGGAGTTTAAAGACCAAGCCCTGGCCAAACATGATGTAACCCCGTCT 15657  
 Db |||||  
 QY 2508 CTACTGAGAT--ATAAATTTAGCTGGGGTGTAGTGCATATGCTGTAAATCCAGCTAGT 2565  
 Db |||||  
 QY 15656 CTACTAAAAATAGAAAAAATTTAGCAGCCATGTTGCGAGCGCCTGTAAATCCAGCTACT 15597  
 Db |||||  
 QY 2566 CGGGAAGCTGAGCGAGGGAATCGCTTGAACCCGAGAGCAGAGGTTGCGGTGAGCTGAG 2625  
 Db |||||  
 QY 15596 TGGGAGGCTGAGCGAGGAGAAATTCCTTGAACCCAGGAAGCAGAGGTTGCGAGTGAGCCGAG 15537  
 Db |||||  
 QY 2626 ATCGTGCCATTTGACTCCAGCCTGGTGACAA-AGCAAGACCTCTGCTCAATAATATAA 2684  
 Db |||||  
 QY 15536 ATCGGCGCACTGCACTCCAGATGGGCAACAAGAGCAAAATCCGCTCTCAAAATAATAA 15477  
 Db |||||

RESULT 14  
 ADB72626/c  
 ID ADB72626 standard; DNA; 96594 BP.  
 XX  
 AC ADB72626;  
 XX  
 DT 04-DEC-2003 (first entry)  
 XX  
 DE Human BLM gene.  
 XX  
 KW human; ds; cytostatic; gene therapy; vaccine; carcinoma; lymphomas;  
 KW cancer; neoplasm; adenocarcinoma; sarcoma; gene.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO2003008583-A2.  
 XX  
 PD 30-JAN-2003.  
 XX  
 PF 26-DEC-2001; 2001WO-US051291.  
 XX  
 PR 02-MAR-2001; 2001US-00798586.  
 PR 23-OCT-2001; 2001US-00004113.  
 PR 08-NOV-2001; 2001US-00052482.  
 PR 30-NOV-2001; 2001US-00997722.  
 PR 20-DEC-2001; 2001US-00034650.  
 XX  
 PA (SAGR-) SAGRES DISCOVERY.  
 XX  
 PI Morris DW, Engelhard EK;  
 XX  
 WPI; 2003-239337/23.  
 DR  
 XX New recombinant nucleic acid, useful for treating carcinomas, lymphomas,  
 PT cancers, neoplasm, adenocarcinoma, or sarcomas.  
 XX  
 PS Claim 1; SEQ ID NO 454; 2304pp; English.  
 XX  
 CC The invention relates to a novel recombinant nucleic acid comprising a  
 nucleotide sequence selected from any of the 660 sequences fully defined

CC in the specification. A polynucleotide of the invention has cytostatic  
CC activity, and may have a use in gene therapy, or in a vaccine. The  
CC recombinant nucleic acids and polypeptides are useful for treating  
CC carcinomas, e.g. lymphomas, cancers, neoplasm, adenocarcinoma, and  
CC sarcomas. The present sequence represents a human gene of the invention.

```
XX SQ Sequence 96594 BP; 27725 A; 19526 C; 19631 G; 29712 T; 0 U; 0 Other;

Query Match      8.2%; Score 245.6; DB 9; Length 96594;
Best Local Similarity 48.2%; Pred. No. 2.8e-53;
Matches 1158; Conservative 0; Mismatches 1174; Indels 68; Gaps 14;

QY 314 GCCGGTGCAGTGGCTAGTCCCTGAAATCCCAACACTTTGGGAGCGGAGGTGGGTGGAT 373
DB 17837 GCCGGGCGTGGTGGCTTATGCCGTGTAATCCAGCACTTTGGGAGCGGAGCGGACAGAT 17778
QY 374 CACTTGAGGTGAGGAGTTGAGACACAGCTGCCCAACATGFTGTAACCCCTGCTCTACTA 433
DB 17777 CACCTGAGGTGAGGAGTTGAGACCATCTGGACCAACATGGCAAAACGCTCTCTACTA 17718
QY 434 AAAATATAAAATTAAGCTGGGCATGGTGGTACCTGTATCCAGATACTTGGGAGG 493
DB 17717 AAAATACAAAATTAAGCAGGCATGATGGTGACACCTTAAACCCAGCTATTACGTAGG 17658
QY 494 TTGAGGCGAGGAATCGCTTGAACCCGGGAGGACAGATGTTGAGTGAACCAAGACTGTGC 553
DB 17657 CTGAGACATGAAATCTCTTGAACCCAGCTGGTAGAGTTGCAGTGAGCCCAAGATCATAC 17598
QY 554 CACTGCATCCAGTCTGGGCAACAGAGTGAGCTTCATCTCAAAACAAACAAACAAAGC 613
DB 17597 CACTGCATCCAGCTGGGCAACAGAGTGAGACTCTTCTCAAAACAAACAAACAAACAA 17541
QY 614 AGTGCCCATCATGTAGGATTCAGTATGATGATGATGATGATGATGATGATGATGATGAT 673
DB 17540 -----CTTCTCTGGGTAGACAAAGGGCTGGGGAGTGAATCAATCAACCAATGGCCAA 17487
QY 674 CTCATAATC-ACCAGTGTGTAGTATCAGTATGATGATGATGATGATGATGATGATGATGAT 732
DB 17486 TTATTTAATCAATCATGTCTAGGTAAAGGAGCCCTACGTAAACCCAGAGGATAGGCT 17427
QY 733 GAGGGTTCAGAAAGATGCCGAGGCGCTTCAAGGTGCTGGGGATGTTGGTGGCAAGCCCTC 792
DB 17426 CAGGGAGCTTCTAGGTTGGTGAGAACATGTAGAGATGGTGGGAGGCTGTTGCTCGCTGG 17367
QY 793 GAATAATAGAAACAGTCTCTGTATTAACAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 852
DB 17366 AGGCATGGAAGTGGCAACCCCTTCCCAATACCTTGTCTGTGTATTTCTTCCAGCTG 17307
QY 853 CCAGGAACCTCAGTAGTAACCTAAGACAGCACCGGTGCTGCTTCCCGAGCGCACCTAGGCCA 912
DB 17306 GCTGCTTCTGAGTTGTTTATAAATAGCTGATAATCTAGTAGGAACTGTTCTGTGACTG 17247
QY 913 GTGGGAAACAGACTACACACAGTCCAGCCAGAGTGCTGAGGCGCAAGATGGGGAA 972
DB 17246 AATGTGAGAGACTTACCACCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 17187
QY 973 GCACGGGAGAGAGGTTCAGGTGGGATGGGAGGGGTTCAGGCAAGAGGGGTTCAGGGCCA 1032
DB 17186 GACAGGAGAGACTTAACATTTGGGGTCTGCATTAACCTGGCAGTTAGTGTTCAGACTG 17127
QY 1033 GGCTGAGGAGAGCCCTGGGACTGTAGGAATTTAGAGGAGGTACCTGACCCCGGATGTTG 1092
DB 17126 AGCTAAATGTAGGACACCCAGTGTCTGCTAAGAACTGAACATATTGTGGGAA----- 17073
QY 1093 GTGAGGAGATTCAGGAAGTCTTCCCTGGAAGAGAGGCTGTGGGAGCTGAGACTCATAAGA 1152
DB 17072 ----AAAAACACCAACATTTGGTGACAGAAAGTGAAGAAATTATTATGGGTACAGAGTAG 17016
QY 1153 TGAGTGGGAGGTTTCCAGGGAGAAAGACCCAGCACCTCAAAAGCATGACTTTTGAGAG 1212
DB 17015 AGAGGAAATGGAGTTTCTTTTACAGTGGTATACGCA-----TAGAATG 16973
QY 1213 AAGCATTCATCCATCAACTGATGAATTTTTCAGACTGGGACGCTGCTCATGCTGTAA 1272
```

```
DB 16972 AAATATTGAGCTTTAAAGAAAAAAATTT--GACTAGATGAAGTGGCTCATGCTTGTAA 16915
QY 1273 TCCAGCACTTTTGAAGGCTGAATGGGAGGATGACTTGGAGCTTAGGCAATTTGTGCAAG 1332
DB 16914 TCCAGCACTTTTGGAGGCCAAGGAGGAGTATCGCTTGGAGCCAGGAGTTTGGAGACCAG 16855
QY 1333 CTGGGCAACATGTTGAGACCC--TGCTTCCAAAAACAAACAAACAAACAAACAAATCAT 1390
DB 16854 CCTAGGCAACAAAGCAAGACCCAGTATCTACAAAAATAAAAAAGTTAGCCAAAGCATGCT 16795
QY 1391 TATACCTCGTACCTACCTGGGTACAGGTAATGAATGACTCAGGCAGATATGCTGCTC 1450
DB 16794 GGTACACTCTGTGTCTCAGCTTCTCGGAGGCTGAGGCAGAGGATCATCTGAGCCCA 16735
QY 1451 TCCTACTTGGGAGAGGGGCTTATCTGCACTAAGCAATAGAGGAGGCAATATATAT 1510
DB 16734 GGAGTTAGAGGCTGCAGTGAGCAACGATCATCTACTGCACTCCAGCCCTGGATAGCTGAG 16675
QY 1511 CCTAAATGAGAGGTACAGATTTGAGACAAACACAGGGCACAGCATATGTACAGGGT 1570
DB 16674 CAAGACCTCTGTGAAAAATAATAAAATACAAATTAAGTTAAATAATAAAAAATTA 16615
QY 1571 AAAGAGGGAATCAGGAGGCTTCTCAGAGAGGTGACATTTAAGCC-----GGACATG 1625
DB 16614 AAAAAAATCTCTGCAACATGAGATGACACAGATGAACCTTGAAGCCATTAAAGCAAGTG 16555
QY 1626 AAGGATGAACGAGTTAGTTTACCAAGGATGGGATGGAAGGGGTGAGAGTGTATGAGGCA 1685
DB 16554 AAGTAAGCCAGTCAAGGACAAATACATGATTTCCACTGCACGAGGATCTTAAATA 16495
QY 1686 GAGGAATCGAGGATCATAGCCCTAGACGGGATCTGACGCCCTTGGAGAAAGTGAGA 1745
DB 16494 GGCACCACTCATAGAAGCAAGAGGGGCACAGTGGTTGCCAGGTAGGGGCTTTCAGGAAA 16435
QY 1746 GAACACAGCCAGTCTGTAGTGGTAAAGTAACAAAGCTGAGAGCCAGGCAATTCCTGT 1805
DB 16434 TGGGGAGTTGTTAATCAAAAAGTATAAAGTTTCAGTAAATGCAAAATGAACCATTTCTAGA 16375
QY 1806 GTCATGCGAGGCTGTGAGTCACGTACAGAGTGTGTTGGGCTTTTGGTTTCTCGGAGCAG 1865
DB 16374 GATCTGCCATACAACTTTGTGCTGCTATACATGTAACACTCTGTTAAGAGG 16316
QY 1866 TCGATTTTAAAGCAGGAAACAGCTGTATTCAGAGTTGGGAAGATCCTGTGTTGCTGCTG 1925
DB 16315 TTAAGAAAAACAATTTAAAAAAATTTTTTTTAGAGACAAGTCTGTGCTCTCACTCAG 16256
QY 1926 AAGGGATGAACCTGGAGGCTTAGAGCCAGGCTGATAGGAGGATCCAGGCTGATGGG 1985
DB 16255 GCTAGCGCAGTGGCAATCATAGTTCACTGCAGCTCAGACTCCTAAGCTCAAGTGA 16196
QY 1986 AGGCTGGAGGTCCGCGGTGATGGACCCAGG---GCTGGGGCCAGGGATGGGAGGAGG 2042
DB 16195 TCCTCTGCTCAGCCTCTGAGAACTTGGAACTACAGGTGCAATCATGCTCACT 16136
QY 2043 AGTAATTTGGAGAGCCCTGGGGCTCTGGCCGAGGAATG-----GATGGTGGGTGAAAC 2096
DB 16135 AATTTTAAAAATTTTATAGAGACAAGTCTCCCAATTTGTTGCCAGACTGCTCTTGAAC 16076
QY 2097 AGGAGAGGAGAGATGCTTAGGCCACTTTGGACACAGTAGGAGGCAAGGACAGGACACC 2156
DB 16075 CTGGCTCAAGAGATCTTGTCTCAGCCTCCCAAAGTCTGGGATTCAGAGAGTGAGC 16016
QY 2157 CAAGGGGAGAGTCCCAAGAGACCAAGCAGGCTGGCAATTCGACAGGGAAGTCTGTCTGG 2216
DB 16015 TACCATGCCAGCCAAAATTTACTTTTAAAAAAAATAAAAAAGAAAGAAACTATATCT 15956
QY 2217 AGCAGGTCTCTGGATAAGGGAGGAAAATGGTGC-----AGTTCCATCTCTCTCCC 2267
DB 15955 ACAACGTGGAAGAATCTCAAAACATAAATAACCAAACTAGAAAGCTCAGTAGAATGCA 15896
QY 2268 TCTCTGTTCAACCTCTAAACTACATGGGGCAGGACCCAGTGGGACTCCATTAATGATG 2327
```

Db 15895 TATAGTATTATCCATTTACAGAGCTCAAAACAGGCAAACTGAGCTATATTAT- 15837  
 QY 2328 GGATGGGTGGATGGAAGGAAGGAGGAAACAACTTTCATTCATCTGGTATTATAC 2387  
 Db 15836 AGATGTATAGATCAGGGCAAACTATCAACAGCAAGAAAGCAATATCACAAAAATTGGGAC 15777  
 QY 2388 AGAACAGGCGAGGTGGGTGCTCAGCCTTGCCATTCTAGCACTTTGGGAGGCTGAGGTGG 2447  
 Db 15776 AATAGGGCCAGGAGCGGTGGCTCATGCCCTGTAAATCCAGCACTTTGGGAGGCGGAG 15717  
 QY 2448 GTGATTACCTCAGGTCAGGAGTTCAAGACCCAGCTTAGCAACGTAAGAGAAACCCCATCT 2507  
 Db 15716 GCAGATTACCTGAGTCAAGAGTTTAAGACCCAGCTTGCCACATGATGTAACCCGCTCT 15657  
 QY 2508 CTACTGAAGAT--ATAAAATTAGCTGGCGTGTAGTGCCATATGCTGTAAATCCAGTAGT 2565  
 Db 15656 CTACTAAAAATAGAAAAAATTAGCCAGCCATGGTGGCGGCGCTGTAAATCCAGTACT 15597  
 QY 2566 CGGGAGCTGAGCGAGGAGTCTTGAACCCGAGGAGGAGTTGGGTGAGCTGAG 2625  
 Db 15596 TGGGAGGCTGAGCGAGGAGTCTTGAACCCGAGGAGGAGTTGGGTGAGCTGAG 15537  
 QY 2626 ATCGTCCACTTGCACTCCAGCTGGGTGACAA-AGCAAGACCTCTGCTCAATTAATAA 2684  
 Db 15536 ATCGGCCACTGCATCCAGATGGGCAACAGCAAGCAAACTCCGCTCAAAATAATAA 15477

RESULT 15  
 AAD41740/c  
 ID AAD41740 standard; DNA; 99500 BP.  
 XX AAD41740;  
 XX 30-OCT-2002 (first entry)  
 XX Human RECQL2 DNA #1.  
 DE Antisense; RECQL2; Bloom's disorder; prophylaxis; infection; tumour;  
 KW inflammation; therapy; human; db.  
 XX Homo sapiens.  
 XX US6399378-B1.  
 XX 04-JUN-2002.  
 XX 01-MAR-2001; 2001US-00798096.  
 XX 01-MAR-2001; 2001US-00798096.  
 XX (ISIS-) ISIS PHARM INC.  
 XX Ward DT, Watt AT;  
 XX WPI; 2002-535979/57.  
 XX Antisense compounds targeted to nucleic acids encoding RECQL2 associated  
 PT with Bloom's disorder, for modulating RECQL2 expression and treating  
 PT diseases e.g. tumors associated with expression of the RECQL2 in humans.  
 XX Example 15; Col 59-144; 86pp; English.  
 XX The invention relates to antisense compounds targetted to nucleic acid  
 CC encoding RECQL2 (gene associated with Bloom's disorder) to inhibit the  
 CC expression of RECQL2. Antisense compounds of the invention are useful for  
 CC treating diseases associated with expression of RECQL2, in humans. They  
 CC are useful for diagnostics, therapeutics and as research reagent, e.g.  
 CC prophylactically to prevent or delay infection, inflammation or tumour  
 CC formation. They are also useful in antisense therapy. The present  
 CC sequence is human RECQL2 DNA  
 XX Sequence 99500 BP; 28415 A; 20198 C; 20481 G; 30406 T; 0 U; 0 Other;  
 XX

Query Match 8.2%; Score 245.6; DB 6; Length 99500;  
 Best Local Similarity 48.2%; Pred. No. 2.8e-53;  
 Matches 1158; Conservative 0; Mismatches 1174; Indels 68; Gaps 14;  
 QY 314 GCCGGTGCAGTGGCTAGTGCCTGAAATCCCAACACTTTGGGAGGCGGAGGTGGGTGGAT 373  
 Db 8960 GCCGGCGTGGTGGCTTATGCTCTGTAATCCCACTTTTGGGAGGCGGAGCGGAGAT 8901  
 QY 374 CACTTGAGTCAAGGATTCGAGACAGCTGCGCAACATGGTGAACCTGTCTCTACTA 433  
 Db 8900 CACCTGAGTCAAGGATTTGAGACCATCTCGCAACATGGCAAAAGCTCTCTACTA 8841  
 QY 434 AAAATATAAAATTTAGCTGGGATGCTGCTGCTTAAATCCAGATCTTTGGGAGG 493  
 Db 8840 AAAATACAAAATTTAGCCAGGATGATGCTGACACCTATAACCCAGCTATTCACTAGG 8781  
 QY 494 TTGAGCAGGAGAAATCGCTTTGAAACCGGAGGAGGATGTTGAGTGAAACCAACACTGTGC 553  
 Db 8780 CTGAGACATGAGAAATCTCTTGAACCCAGCTGGTAGAGTTGCACTGAGCCAGATCATAC 8721  
 QY 554 CACTGCACCTCCAGTCTGGGCAACAGAGTGAGCTCCATCTCAAAACAAACAAACAAAGC 613  
 Db 8720 CACTGCACCTCCAGCTGGGCGACAGAGTGAGACTCTCTCAAAACAAACAAACAA-- 8664  
 QY 614 AGTGCCCATCATGTAGGATTTGAGTGTGAGTGAGGACTGAGCTTTGTGCAAAAGTGAGCA 673  
 Db 8663 -----CTTCTCTGGTAGGACAAAGGGCTGGGGAGTGACTTAATCAACCAATGGCCAA 8610  
 QY 674 CTCACATAATC-ACCAGTTGTAGTATCAGTGATTAACATCAATGATCCAGTAAAGCCCT 732  
 Db 8609 TTATTTAATCAATCATGTCTAGGTAAGAGGCGCTACGTAAAAAACCAGAGGATAGGGCT 8550  
 QY 733 GAGGTTTCAAGAGATGCGGAGCGCTTTCAAGGTGCTGGGGATTTGGTGGGCAAGCCCTC 792  
 Db 8549 CAGGAGCTTCTAGGTTGGTGAGAACATGTAGAGATGGTGGGAGGCTGGTGTGCTCGGAG 8490  
 QY 793 GAATAATAGAAAACAGTTCTCTGTATTACAAAGAGCAGAGAGGCCCATGCTGGGTGCTG 852  
 Db 8489 AGGGCATGGAAGTGCCACACCCCTTCCACATACCTTGTCTGTATTTCCTCCAGCTG 8430  
 QY 853 CCAGGAATCTCAGTAGTAACTAAGACAGCACCGGTGCTGCTTCCCGAGCGCACTAGGCCA 912  
 Db 8429 GCTGCTTCTGAGTTGTTTATAATAGCTGATAATCTAGTAAGAAAACCTGTTCTGTGACTG 8370  
 QY 913 GTGGGAAACAGACTCACACACAGTCCAGGCCAGAGTGCTGAGGCGCCAGATGGGGA 972  
 Db 8369 AATGTGAGGACTTACCATTTGCTGTGTAGTCTGAAGTGGGGGTCTGGAATCGGGGTG 8310  
 QY 973 GCACGGGAGAAAGGTCAAGGTGGGATGGGAGGGGTCAAGGCAAGAGGGGTCAAGGCCA 1032  
 Db 8309 GACAGGGAGACTTAACATTTGGGGTCTGCACTAAACACCTGGCAGTTAGTGTCAAGACCTG 8250  
 QY 1033 GGCTGAGGGAAGCCCTGGGACTGTAGGAATTTAGAGGAGGTACTGACCCGGCATGTTG 1092  
 Db 8249 AGCTAAATGTAGGACACCCAGTGTCTGCTAAAGAACTGAACTATTGTTGGGGA-- 8196  
 QY 1093 GTGAGGAGATTCAGGAAGTCTTCTTGGAAAGAGAGGCTGTCGGAGCTGAGACTCATAAGA 1152  
 Db 8195 ----AAACACACACATTTGGTGACAGAAGTGAAGTAATTATATATGGGTACAGACTAG 8139  
 QY 1153 TGAGTGGGAGGGTGTTCAGGCGAGAAAGACCACTCAACAAAGCATGACTTTTGAGAG 1212  
 Db 8138 AGAGGAATCGAGTTTCTTTTACAGTGGTATACGCA-----TAGAATG 8096  
 QY 1213 AGCATTTCATTCATTAACCTGATGAATTTTCAAGCTGGGCAAGCTGGCTCATGCTCTATA 1272  
 Db 8095 AAATATTTCAGCTTTAAAAAGAAAAAATT--GACTAGATGAAGTGGCTCATGCTTTGTA 8038  
 QY 1273 TCCAGCACCTTTTGAAGGCTGAATGGGAGGATGACTTTGAGCCCTAGGCACTTTGTGACAAG 1332  
 Db 8037 TCCAGCACCTTTGGAGGCCAAGGACAGGATGATGCTTTGAGCCCGAGGATTTGAGACAG 7978  
 QY 1333 CCTGGCAACATGTTGAGACCC---TGCTCCACAAACAAACAAACAAACAAATCAT 1390

Db 7977 CCTAGGCAACAAGCAAGACCCAGTATCTACAAAAAATAAAGATTACCCAGCATGCT 7918  
QY 1391 TATACCTGGTACCATGGGTACCAAGGTACATAGAAATGACTCAGGCAGATATGGTGTCTTC 1450  
Db 7917 GGTAACACTCTGTGTCTCACCCTTCGCGGAGGCTGAGGAGAAGGATCATCTGAGCCCA 7858  
QY 1451 TCCTACTGTGGAGAGGGGGCTTATCTGCTAGTAAGACAATAGAGGGAGGAATATAAT 1510  
Db 7857 GGAGTTAGAGGCTGCAGTGAACAACGATCATCTACTGCACTCCAGCCTGGATAGCTGAG 7798  
QY 1511 CCTAAATGAGAGGTACAGATTGTAGAGCAAAACACAGGCAACAGGCATATGTACAGGGT 1570  
Db 7797 CAAGACCCCTGTCTGAAAAAATAAATAACATACAATTAAGCTTAATAATAAATAAATA 7738  
QY 1571 AAAGAGGAATCAGGGAAGGCTTCTCAGAGAAGGTGACATTTAAGCC-----GGACATG 1625  
Db 7737 AAAAAAATCCTGCACATGAGATGACACAGATGAACCTTTGAAGCCATTGAAGCAAAAGTG 7678  
QY 1626 AAGGATGAACGAGTTAGTTTACCAAGGATGGATGGAAGGGGTGAGAGTGTGAGGCA 1685  
Db 7677 AAGTAAGCCAGTTCACAGGACAAATACTACATGATTTCCACTGCAAGCAGGTATCTAAATA 7618  
QY 1686 GAGGAACTGACGATCATAGCCCTAGACAGGGATCCTGACGCCCTTTGAGGAAGTGAGA 1745  
Db 7617 GGCAACTCATAGAAGCAAGAGGGGCACAGTGGTTCAGGGTAGGGCTTCAGGAAA 7558  
QY 1746 GAAGACCAGCCAGTCTGTAGTGGGTAACTAAGTAACAAGCTGAGAAGCCAGGAAATCCCTG 1805  
Db 7557 TGGGAGTGTATTAACAAAAGTAAGATTTTCCAGTAAAGTTTCCAGTAAAGCAATGAACATTTCTAGA 7498  
QY 1806 GTCAATGAGGCGCTGTGAGTCACTCAGAGTGTGTGGGCTTTTGTCTTCTCGGGAGCAG 1865  
Db 7497 GATCTGCCATACAACATTTGGC--TGTTGTGTACTATACATGTAATACTCTGTTAAGAGG 7439  
QY 1866 TCGATTTTACGAGGAAACAGCTGTATTCAGAGTTGGGAAGATCCTGTGGTTGCTGCTG 1925  
Db 7438 TTAAGAAAAACAATTTAAAAAAATTTTATAGAGCAAGGTCTTGTCTGTCACTCAG 7379  
QY 1926 AAGGGGATGAACCTGGAGGCTAGGAGCCAGGCTGATAGGAGGATCCAGGGTATGGGG 1985  
Db 7378 GCTAGAGCGAGTGGCAACATCATAGTTCACCTGAGCCTCAGACTCCAAAGCTCAAGTGA 7319  
QY 1986 AGGCTGGAGGTCGCGGTGATGACACAGG---GCTGGGGCCAGGGGATGGGAGGAGG 2042  
Db 7318 TCCTCTGCTCAGCCTCTGAGAACTCTGGAACCTACAGGTGCAATCATCTCACT 7259  
QY 2043 AGTAATTTGGAGAGGCGTGGGGCTCTGGCCGAGGAATG-----GATGGTGGGCTGAAAC 2096  
Db 7258 AATTTTAAATTTTATAGAGACAAGTCTCCCATTTGTCAGAGCTGGTCTTGAACCTC 7199  
QY 2097 AGGAGAGGAGAGATGCTTAGGCCACTTTTGGACACAGTAGGGCAAGGACAGGAGACAC 2156  
Db 7198 CTGGCCTCAGAGATCTTGTCTCCTCAGCCTCCAAAGTGTGGGATACAGAAAGTGAGC 7139  
QY 2157 CAAGGGGAAGTCCCAAGAGACACACAGCAGGCTGGCATTTGGACAGGAGGCTGTCTGG 2216  
Db 7138 TACCATGCCAGCAAAATTTACTATTAAAAAAGAAAAAGAAAAAGAACTATACT 7079  
QY 2217 AGCAGTGTCTTGGATAAGGAGGAAAAATGGTGC-----AGTTCCATCTCCTCC 2267  
Db 7078 ACACGTGGAAGAATCTCAACAACATAAATACCAAACTAGAAAGACTCAGTAGAATGCA 7019  
QY 2268 TCTCTCTTCAACTCTAAACTACATGGGCAAGGACCCAGGAGTGGGACTCCATAAATGATG 2327  
Db 7018 TATAGTATTATCCATTTACAGGCTCAAAACAGGCAAACTGAGCTATATTATTAT- 6960  
QY 2328 GGATGGTGGATGGAAGGAGGAGGAGAAACAATCTTCAATTCATCTGTTATTAC 2387  
Db 6959 AGATGTATAGATCAGGGCAAACTATGAACAAGAAAGCAATATCAAAAAATTTGGGAC 6900  
QY 2388 AGAACAGGCCAGGTGGGTGCTCAGCTTGGCCATTTAGCACTTTGGGAGGCTGAGGTGG 2447

Db 6899 AATAGGGCCAGAGCGGTGGCTCATGCTCTGTAAATCCAGCACTTTGGGAGGCCAGGAG 6840  
QY 2448 GTGGATTACCTCAGGTCAGGAGTTCAAGACCAAGCCTTAGACAACGTTAGAGAAAACCCCATCT 2507  
Db 6839 GCAGATTACCTCAGGTCAGGAGTTTAAAGACCAAGCCTTGGCCAAACATGATGTAACCCGTCT 6780  
QY 2508 CTACTGAAGAT--ATAAAAATTAGCTGGGCGTGTAGTGGCATATGCCCTGTATCCAGCTACT 2565  
Db 6779 CTACTAAAAATAGAAAAAATTAGCCAGCCTAGTGGCAGGCGCTGTAAATCCAGCTACT 6720  
QY 2566 CGGGAAGCTGAGGCAAGGAGATCGCTTGAACCCAGAGAGGAGGTTGGGTTGAGTGTAG 2625  
Db 6719 TGGGAGGCTGAGGCGAGGAGAAATTGCTTGAACCCAGCAAGCAGAGGTTGCACTGAGCCGAG 6660  
QY 2626 ATCGTGCAATTGCACTCCAGGCTGGGTGACAA--AGCAAGACCTCGTCTCAATAATAATAA 2684  
Db 6659 ATCGGCCACTGCACTCCAGAAATGGGCAACAAGACAAAACCTCCGCTCTCAAAATAATAA 6600

Search completed: September 17, 2004, 08:42:04  
Job time : 1143.52 secs

***This Page Blank (uspto)***



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 17, 2004, 07:38:50 ; Search time 192.326 Seconds

(without alignment)  
8659.273 Million cell updates/sec

Title: US-10-669-693-3\_COPY\_1980\_4980

Perfect score: 3001

Sequence: 1 cttcttcctccctaccoccc.....atcctggtaacacagtgaa 3001

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents NA:\*
- 1: /cgn2\_6/ptodata/2/ina/5A COMB.seq:\*
  - 2: /cgn2\_6/ptodata/2/ina/5B COMB.seq:\*
  - 3: /cgn2\_6/ptodata/2/ina/6A COMB.seq:\*
  - 4: /cgn2\_6/ptodata/2/ina/6B COMB.seq:\*
  - 5: /cgn2\_6/ptodata/2/ina/PTUS COMB.seq:\*
  - 6: /cgn2\_6/ptodata/2/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	245.8	8.2	786431	US-09-751-389-3	Sequence 3, Appli
C 2	245.6	8.2	99500	US-09-798-096-10	Sequence 10, Appl
C 3	240.8	8.0	17327	US-07-906-871-15	Sequence 15, Appl
C 4	234	7.8	21968	US-09-851-985-3	Sequence 3, Appli
C 5	232.6	7.8	2839	US-09-061-702-1	Sequence 1, Appli
C 6	232.2	7.7	48763	US-09-916-204-3	Sequence 3, Appli
C 7	232	7.7	66804	US-09-740-041-3	Sequence 3, Appli
C 8	231.4	7.7	72928	US-09-009-913-1	Sequence 1, Appli
C 9	231.2	7.7	128779	US-09-497-855A-38	Sequence 38, Appl
C 10	230.6	7.7	174493	US-09-804-471A-3	Sequence 3, Appli
C 11	230.6	7.7	174493	US-10-238-709-3	Sequence 3, Appli
C 12	230.4	7.7	16063	US-09-021-052-3	Sequence 3, Appli
C 13	230.4	7.7	16063	US-10-020-121-3	Sequence 3, Appli
C 14	229	7.6	246240	US-08-724-394A-20	Sequence 20, Appl
C 15	229	7.6	246240	US-08-724-394A-21	Sequence 21, Appl
C 16	229	7.6	246240	US-08-724-394A-22	Sequence 22, Appl
C 17	228.8	7.6	90541	US-09-759-359A-3	Sequence 3, Appli
C 18	228.4	7.6	11288	US-08-646-301A-1	Sequence 1, Appli
C 19	228.4	7.6	11288	US-08-481-968A-4	Sequence 4, Appli
C 20	228.4	7.6	11288	US-08-154-712B-4	Sequence 4, Appli
C 21	228.4	7.6	15056	US-09-474-699-10	Sequence 10, Appl
C 22	228.2	7.6	685	US-09-183-266A-16	Sequence 16, Appl
C 23	228.2	7.6	53332	US-09-801-861-3	Sequence 3, Appli
C 24	228	7.6	116592	US-09-818-512-3	Sequence 3, Appli
C 25	227.6	7.6	45716	US-08-965-048-5	Sequence 5, Appli
C 26	227.6	7.6	45989	US-08-965-048-6	Sequence 6, Appli
C 27	227.6	7.6	128779	US-09-497-855A-38	Sequence 38, Appl

C 28	227.2	7.6	36651	4	US-09-738-894A-3	Sequence 3, Appli
C 29	227.2	7.6	36651	4	US-09-964-469-3	Sequence 3, Appli
C 30	227	7.6	36741	3	US-09-301-665-3	Sequence 3, Appli
C 31	226.6	7.6	70000	4	US-09-851-896-3	Sequence 3, Appli
C 32	226.2	7.5	399	4	US-09-621-976-13959	Sequence 13959, A
C 33	226.2	7.5	10380	3	US-09-077-354B-3	Sequence 3, Appli
C 34	226	7.5	497	4	US-09-621-976-3876	Sequence 3876, Ap
C 35	226	7.5	3627	4	US-09-323-873A-6	Sequence 6, Appli
C 36	226	7.5	99500	4	US-09-798-096-10	Sequence 10, Appli
C 37	225.8	7.5	36651	4	US-09-738-894A-3	Sequence 3, Appli
C 38	225.8	7.5	36651	4	US-09-964-469-3	Sequence 3, Appli
C 39	225.2	7.5	368	4	US-09-621-976-10146	Sequence 10146, A
C 40	225.2	7.5	2115	1	US-08-395-800A-7	Sequence 7, Appli
C 41	225.2	7.5	118067	4	US-09-497-855A-32	Sequence 32, Appl
C 42	224.6	7.5	62804	4	US-09-800-960-3	Sequence 3, Appli
C 43	224.6	7.5	62804	4	US-10-096-960-3	Sequence 3, Appli
C 44	224.4	7.5	283	4	US-08-579-445-26	Sequence 26, Appl
C 45	224.4	7.5	841	5	PCT-US93-06251-80	Sequence 80, Appl

## ALIGNMENTS

RESULT 1  
US-09-751-389-3/c  
; Sequence 3, Application US/09751389  
; Patent No. 6630334  
; GENERAL INFORMATION:  
; APPLICANT: GUEGLER, Karl et al  
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC  
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES  
; TITLE OF INVENTION: THEREOF  
; FILE REFERENCE: CL001067  
; CURRENT APPLICATION NUMBER: US/09/751.389  
; CURRENT FILING DATE: 2001-01-02  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 786431  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)..(786431)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-751-389-3

Query Match	8.2%	Score 245.8;	DB 4;	Length 786431;
Best Local Similarity	84.2%	Pred. No. 7.8e-57;		
Matches 277;	Conservative 0;	Mismatches 52;	Indels 0;	Gaps 0;
QY	282	TCGTCAATGGAGTGATACAGTCAGCGGGTGCAGTGGCTAGTGGCTAGTGGCTGAAT	341	
Db	572672	TTTTTTCATTACTGTATTAACTGACTAGAGCCAGGTCTGTGGCTCACTATAAT	572613	
QY	342	CCCAACACTTTGGAGCGCGGAGTGCGGTGATCACTTGGAGTCCAGGATTCGAGACCAGC	401	
Db	572612	CCGAGCACTTTGGAGCGCGGAGTGCGGTGATCACTTGGAGTTCGAGACCAGC	572553	
QY	402	CTGGCCAAACATGGTGAACCCCTCTCTACTAAAAATATAAAATTAAGCTGGGATGGTG	461	
Db	572552	CTGGCCAAACATGGTGAACCCCTCTCTACTAAAAATATAAAATTAAGCTGGGATGGTG	572493	
QY	462	GTGCGTACCTGTAATCCAGATACCTTGGAGGTTGAGGCGAGGAGATCGCTGAACCCGG	521	
Db	572492	GGGTACCTGTGATCCAGCTACTTTGGAGGTTGAGGCGAGGAGATCACTTGAACCTGG	572433	
QY	522	GAGGCGAGATGTTGCAGTGAACCAAGACTGTGCACCTGCACCTCCAGTCTGGGCAACAGAGT	581	
Db	572432	GAGGTGGAGTTGCACTGAGTCAAGATCGTGCACCTGCACCTCCAGTCTGGGCAACAGAGT	572373	
QY	582	GAGCCTCCATCTCAACAAACAAACAAA	610	

Db	572372	GAAACTCGTCTCAAAACAACAACAAA 572344	
RESULT 2			
US-09-798-096-10/c			
; Sequence 10, Application US/09798096			
; Patent No. 6399378			
; GENERAL INFORMATION:			
; APPLICANT: Donna T. Ward			
; APPLICANT: Andrew T. Watt			
; TITLE OF INVENTION: ANTISENSE MODULATION OF RECQL2 EXPRESSION			
; FILE REFERENCE: RTS-0207			
; CURRENT APPLICATION NUMBER: US/09/798,096			
; CURRENT FILING DATE: 2001-03-01			
; NUMBER OF SEQ ID NOS: 89			
; SEQ ID NO 10			
; LENGTH: 99500			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
; FEATURE:			
US-09-798-096-10			
Query Match			
Best Local Similarity 8.2%; Score 245.6; DB 4; Length 99500;			
Matches 1158; Conservative 0; Mismatches 1174; Indels 68; Gaps 14;			
QY	314	GCCTGAGGAGCCCTCGGACTGTAGGAATTTAGAGGAGGTACCTGACCCCGCATTTTG 1092	
Db	8249	AGCTAAATGTAGGACACCCAGTGTCTGCTAAAGAACTGAACTATTGTGGGAAA----- 8196	
QY	1093	GTGAGGGAGATTTCAGGAAGTCTTCCTGGAAGAGGCTGTGCGAGCTGAGACTCATAGA 1152	
Db	8195	---AAAACACCAACACATTTGGTGACAGAAGTGAGGAATTTATATGGGTACAGAGTAG 8139	
QY	1153	TGAGTGGGAGGGTGTTCAGGCAGAAACACGACACCTACAAAAGCATGACTTTGAGAG 1212	
Db	8138	AGAGGAATGGAGTTTCTTTTACAGTGATACGCA-----TAGAATG 8096	
QY	1213	AAGCATTCATCCATTCACATGATGAATTTTCAGACTGGGCACGCTGGGTCAATGCTGTAA 1272	
Db	8095	AAATATTGAGCTTTTAAAGAAAGAAATTT--GACTAGATGAAGTGGCTCATGCTGTAA 8038	
QY	1273	TCCAGCACTTTGGAAGCTGAATGGGAGAGATGACTTGAGCCTAGGCATTTGTGACAAAG 1332	
Db	8037	TCCAGCACTTTGGGAGGCCAAGGCGAGGATATCGCTTGAGCCCGAGGATTTGAGACCCAG 7978	
QY	1333	CCTGGGCAACATGGTGAGACCC--TGCCTCCACAAAACAAAACAAAACAAAATCAT 1390	
Db	7977	CCTAGGCAACAAAGCAAGACCCAGTATCTACAAAATAAATAAGTTAGCCAGCATGGT 7918	
QY	1391	TATACCTCGTACCATGGGTACCAGGTACATAGAAATGACTCAGGCAGATATGGTGTCTC 1450	
Db	7917	GGTACACT 7858	
QY	1451	TCCTACTGTGGAGAGCGGGCTTATCTGCTAGTAAAGCAATAGAGGAGGGAATATAAT 1510	
Db	7857	GGAGTTAGAGGCTGCACTGAGCAACGATCATTTCTACTGCACTCCAGCCTCGGATAGT 7798	
QY	1511	CCTAAATGAGAGGTACAGATTTGAGAGCAACACAGGSCACAGGCATATGTACAGGGGT 1570	
Db	7797	CAAGACCCCTCTGAAAAATAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 7738	
QY	1571	AAAGAGGGAATCAGGGAAGGCTTCTCAGAGAAGGTGACATTTTAAGCC-----GGGACATG 1625	
Db	7737	AAAAAAATCTGCAACATGAGATGACACAGATGAACCTTTGAAGCCATAAGCAAGTG 7678	
QY	1626	AAGGATGAACGAGTTAGTTTCCACAGGATGGGATGGAAGGGGTGAGATGAGGAGCA 1685	
Db	7677	AAGTAAGCCAGTCACAGGACAAATACTACATGATCTCCACTGCAACAGAGGTATCTAAATA 7618	
QY	1686	GAGGNACTGCGAGATCATAGGCTAGACAGGGGATCTTGACGCCCTTGAGGAAGTGAGA 1745	
Db	7617	GGCAAACTCATAGAAGCAAGAGGGGCAAGTGGTTGCCAGGGTAGGGGCTTTCAGGAAA 7558	
QY	1746	GAAGACCGCGCAGTCGTAGTGGTTAAGTAACAAAGCTGAGAAGCCAGGGAATCCCTG 1805	
Db	7557	TGGGGAGTTGTTAATCAAAAAGTATAAAGTTTCAGTAATGCAAAATGAACGCATTCTAGA 7498	
QY	1806	GTCAATCGAGGGCTGTGAGTCAGTCAGAGTGTTCGGGCTTTTGTCTCTCTCTCTCTCTCT 1865	
Db	7497	GATCTGCCATACAACTTTGTGTC--TGTGTGTTACTATACATGTAAAACTCTGTTAAGAG 7439	
QY	1866	TCGATTTTAAAGCAGGAGACAGCTGTATTCAGAGTTGCGAAGATCTGTGTTGCTGCTG 1925	
Db	7438	TTAAGAAAACAAATTTAAAAAAATTTTATTTTATAGACAGAGTCTTCTCTCTCTCTCTCT 7379	
QY	1926	AAGGGATGAAACTGGAGGCTAGGAGCCAGGGTGATAGGAGGATCCAGGGTGATGGGG 1985	
Db	7378	GCTAGAGCGCAGTGGCAACATCATAGTTCACTGCAGCCTCAGACTCTCAAGCTCAAGTGA 7319	
QY	1986	AGCTGGGAGGTCGGGTGATGGACGAG--GCTGGGCGCAGGGGATGGGAGGAAGG 2042	
Db	7318	TCCTCTGCTCAGCCTCTGGGAATCTGGAACCTACAGGTGTCATGATCATGCTCACT 7259	
QY	2043	AGTAATTTGGGAGAGGCTGGGGCTCTGGCCGAGGAATG-----GATGTGGGTGGAAC 2096	
Db	7258	AAATTTTAAATTTTATAGACAAAGTCTCCCATTTGTTGCCAGACTGGTCTTGAATCTC 7199	



;; TITLE OF INVENTION: ISOLATED HUMAN RECEPTOR PROTEINS,  
;; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN RECEPTOR PROTEINS, AND  
;; FILE OF INVENTION: USES THEREOF  
;; FILE REFERENCE: CL001238  
;; CURRENT APPLICATION NUMBER: US/09/851,985  
;; CURRENT FILING DATE: 2001-05-10  
;; NUMBER OF SEQ ID NOS: 4  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 3  
;; LENGTH: 21968  
;; TYPE: DNA  
;; ORGANISM: Human  
;; FEATURE:  
;; NAME/KEY: misc\_feature  
;; LOCATION: (1)...(21968)  
;; OTHER INFORMATION: n = A,T,C or G  
US-09-851-985-3

Query Match 7.8%; Score 234; DB 4; Length 21968;  
Best Local Similarity 82.9%; Pred. No. 2.1e-54;  
Matches 267; Conservative 0; Mismatches 55; Indels 0; Gaps 0;  
QY 314 GCGGGTGCAGTGGCTAGTGCCTGAAATCCCAACACTTTGGGAGGCGGAGTGGTGGAT 373  
DB 12119 GCTGGGCGGGTGGCTCAGCTGTATCCAGACTTTGGGAGGCGGAGTGGAT 12178  
QY 374 CACTTGAGTCTAGGAGTTCGAGACCAAGCTGCGCAACATGGTGAAACCCCTGTCTACTA 433  
DB 12179 CAGCTGAGTTCAGGAGCTCAAGACCAAGCTGCGCAACATGGTGAAACCCCGTCTCAGCTG 12238  
QY 434 AAAAAATAAAATAGCTGGGCTAGTGGCTGCTGCTGTAATCCAGATCTTGGGAGG 493  
DB 12239 AAAAAATAAAATAGCTGGGCTAGTGGCTGCTGCTGTAATCCAGATCTTGGGAGG 12298  
QY 494 TTGAGCAGGAGAAATCGCTTGAACCGGAGGAGGAGTGTGAGTGAACCAAGACTGTGC 553  
DB 12299 CTGAGGAGGAGAAATCACTTGAACCGGAGGAGGAGTGTGAGTGAACCAAGACTGTGC 12358  
QY 554 CACTGCACCTCCAGTCTGGGCAACAGAGTGAAGCTTCATCTCAACAAACAAACAAAGC 613  
DB 12359 CACTGCACCTCCAGCTGGGCAACAGAGTGAAGCTTCATCTCAACAAACAAACAAAGC 12418  
QY 614 AGTGGCCATCATGTAGATTGA 635  
DB 12419 ATAGCCAGTTGCTAGATTGA 12440

RESULT 5  
US-09-061-702-1/c  
; Sequence 1, Application US/09061702  
; Patent No. 6165737  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Xiaodong  
; APPLICANT: Liu, Xuecong  
; TITLE OF INVENTION: DNA FRAGMENTATION FACTOR INVOLVED IN  
; TITLE OF INVENTION: APOPTOSIS  
; NUMBER OF SEQUENCES: 21  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Arnold, White & Durkee  
; STREET: P.O. Box 4433  
; CITY: Houston  
; STATE: TX  
; COUNTRY: USA  
; ZIP: 77210-4433  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; APPLICATION NUMBER: US/09/061,702  
; FILING DATE: Concurrently Herewith  
; CLASSIFICATION:

;; ATTORNEY/AGENT INFORMATION:  
;; NAME: McMillian, Nabeela R.  
;; REGISTRATION NUMBER: P-43,363  
;; REFERENCE/DOCKET NUMBER: UTSD:546  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (512) 418-3000  
;; TELEFAX: (512) 474-7577  
;; INFORMATION FOR SEQ ID NO: 1:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 2839 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
US-09-061-702-1

Query Match 7.8%; Score 232.6; DB 3; Length 2839;  
Best Local Similarity 71.9%; Pred. No. 1.7e-54;  
Matches 304; Conservative 0; Mismatches 119; Indels 0; Gaps 0;  
QY 228 CACCACTTACTGCTGTGACCTTTGCACATGGCTTAGTCCCTCTGTTGCCCTCATCTGTC 287  
DB 2361 CAGCAGTTTCAGTGGTTCACGTCATAGCTGTTCAACAGACATTTGGGGGGTTTCTATT 2302  
QY 288 AATGGAGTGATAACAGTGCCTCATCAGCGGGTGCAGTGGCTAGTGCCTGAAATCCCAAC 347  
DB 2301 TCTTCCATTATAACAATTCAGCAGTTCAGTGCAGTGGCTCATGCTGTATATCCAGC 2242  
QY 348 ACTTTGGGAGGCGGAGTGGTGGATCACTTGAAGTCAAGAGTTCGAGACGAGCTGGCC 407  
DB 2241 ACTTTGGGAGGCGGAGTGGTGGATCACTTGAAGTCAAGAGTTCGAGACGAGCTGGCC 2182  
QY 408 AACTGCTGCAACCCCTGCTCTACTTAAATAATAAAATAGTGGGCTAGTGGTGGCT 467  
DB 2181 AACTGCTGCAACCCCTGCTCTACTTAAATAATAAAATAGTGGGCTAGTGGTGGCT 2122  
QY 468 ACTTGTATATCCAGATCACTTGGGAGTTCGAGGAGGAGTTCGAGGAGGAGTTCGAGGAGG 527  
DB 2121 GCTGTATATCCAGATCACTTGGGAGTTCGAGGAGGAGTTCGAGGAGGAGTTCGAGGAGG 2062  
QY 528 GATGTTGAGTGAACCAAGACTGTGCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 587  
DB 2061 GAGGTTGAGTGAACCAAGACTGTGCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2002  
QY 588 CCATCTCAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 647  
DB 2001 CTGCTCAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 1942  
QY 648 GGA 650  
DB 1941 GTA 1939

RESULT 6  
US-09-916-204-3  
; Sequence 3, Application US/09916204  
; Patent No. 6638745  
; GENERAL INFORMATION:  
; APPLICANT: Wei, Ming-Hui et al.  
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC  
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES  
; TITLE OF INVENTION: THEREOF  
; FILE REFERENCE: CL001164CIP  
; CURRENT APPLICATION NUMBER: US/09/916,204  
; CURRENT FILING DATE: 2001-07-24  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 48763  
; TYPE: DNA  
; ORGANISM: Human  
US-09-916-204-3  
Query Match 7.7%; Score 232.2; DB 4; Length 48763;

```

Best Local Similarity 85.7%, Pred. No. 1e-53;
Matches 258; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY 314 GCCGGGTGCAGTGGCTAGTGGCTGAAATCCCAACACATTTGGAGGCGGAGGTGGGTGGAT 373
Db 40920 GCCGGCGCAGTGGCTCACACCTGTATCCAGCACATTTGGAGACGAGGCGGTGGAT 40979

QY 374 CACTTGAGGTCAGAGTTTCGAGACCGACCTGGCCAAATGGTGAACCCCTGTCTCTACTA 433
Db 40980 CACTTGAGGTCAGAGTTTCGAGACTGGCCCTGGCCAAACATGATGAAATCCCGTCTCTACTA 41039

QY 434 AAAATATAAAAAATTAGCTGGGCATGGTGGTGGCTACTGTAAATCCAGATACTTTGGGAGG 493
Db 41040 AAAATACAAAAATTAGCTGGGCATGGTGGCGCATGTCTGTAGTCCAGCCACTTCGGAGG 41099

QY 494 TTGAGGCAGGAGAAATCGCTTTGAACCCGGGAGGCAGATTTGCAGTGAACCAAGACTGTGC 553
Db 41100 CTGAGGCAGGAGAAATCGCTTTGAACCCGGGAGGCAGAGTTGCAGTCAGCCGAGATCGTGG 41159

QY 554 CACTGCACTCCAGTCTGGGCAACAGAGTGGCCCTCCATCTCAACAAACAAACAAAAGC 613
Db 41160 CACTGCACTCCAGCTGGGTAAACAGATGAGACTTCATCTCAAAAAAAACCCCAAAAAAC 41219

QY 614 A 614
Db 41220 A 41220

RESULT 7
US-09-740-041-3/c
; Sequence 3, Application US/09740041
; Patent No. 6562593
; GENERAL INFORMATION:
; APPLICANT: MERKULOV, Karl et al
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL001001
; CURRENT APPLICATION NUMBER: US/09/740,041
; CURRENT FILING DATE: 2000-12-20
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 66804
; TYPE: DNA
; ORGANISM: Human
US-09-740-041-3

```

```

Db          31325  TGTG 31322

      III
RESULT 8
US-09-009-913-1/c
; Sequence 1, Application US/09009913
; Patent No. 6087485
; GENERAL INFORMATION:
;   APPLICANT: AXYS Pharmaceuticals, Inc.
;   TITLE OF INVENTION: Asthma Related Genes
;   NUMBER OF SEQUENCES: 339
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Bozicevic & Reed, LLP
;   STREET: 285 Hamilton Ave, Suite 200
;   CITY: Palo Alto
;   STATE: CA
;   COUNTRY: USA
;   ZIP: 94301
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Diskette
;   COMPUTER: IBM Compatible
;   OPERATING SYSTEM: DOS
;   SOFTWARE: FastSeq for Windows Version 2.0
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/09/009,913
;   FILING DATE: 21-JAN-1998
;   CLASSIFICATION:
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER:
;   FILING DATE:
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Sherwood, Pamela J
;   REGISTRATION NUMBER: 36,677
;   REFERENCE/DOCKET NUMBER: SEQ-4P
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 650-327-3231
;   TELEFAX: 650-327-3231
;   TELEX:
;   INFORMATION FOR SEQ ID NO: 1:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 72928 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: double
;   TOPOLOGY: linear
;   MOLECULE TYPE: Genomic DNA
US-09-009-913-1

```

Query Match	7.7%;	Score 232;	DB 4;	Length 66804;
Best Local Similarity	85.2%;	Pred. No. 1.4e-53;		
Matches 259;	Conservative 0;	Mismatches 45;	Indels 0;	Gaps 0;
QY	314	GC CGGTGCAGTGGCTAGTGCCTGAAATCCCAACACTTTGGGAGCGGAGTGGGTGGAT	373	
Db	31625	GCTGGCGCAGTGGCTCATGCTGTAATCCGACACTTTGGGAGCCCAAGCGCGGTAGAT	31566	
QY	374	CAC TTGAGGTCAGGAGTTCGAGACCAGCCTGGCCACATGGTGAACCCCTGCTCTACTA	433	
Db	31565	CAC TTGAGGTCAGGAGTTCGAGCAGCCTGGCCAAATGTTGTAACCCCGTCTCTACTA	31506	
QY	434	AAAAATAAAAAATTAGCTGGGCATGTTGGTGGCTACCTGTATATCCAGATACTCTGGGAGG	493	
Db	31505	AAAAATAAAAAATTAGCCTGGCATGTTGGCA CAT GCCTGTAAATCCCACTACTCTAGGAGG	31446	
QY	494	TTGAGGCAGAGAAATCGCTTTGAACCCGGGAGGCAGATGTTGCAGTGAACCAAGACTGTGC	553	
Db	31445	CTGGGCGAGGAGAAATCGCTTTGAACCTGGGAGCGGAAGTTGCAGTGAAGCCAGATTGTGC	31386	
QY	554	CAC TGCAC TCCAGTCTGGCCACAGAGTGAAGCCCTCCATCTCAAAACAAACACAAAAGC	613	
Db	31385	CAC TGCAC TCCAGCCTGGGTGACAGAGTAAGACCCCATCTCAAAAAA AAAAAAAGGAAA	31326	
QY	614	AGTG 617		







Db 10992 GAGCGGGTGGATCAGGAGTTCAGGAGATCGAGACCATCTCGCCAAACATGTTGAA 11047

RESULT 14

US-08-724-394A-20/c

; Sequence 20, Application US/08724394A

; Patent No. 5872237

; GENERAL INFORMATION:

; APPLICANT: Feder, John N.

; APPLICANT: Kronmal, Gregory S.

; APPLICANT: Lauer, Peter M.

; APPLICANT: Ruddy, David A.

; APPLICANT: Thomas, Winston

; APPLICANT: Tsuchihashi, Zenta

; APPLICANT: Wolff, Roger K.

; TITLE OF INVENTION: Megabase Transcript Map: No. 5872237el

; TITLE OF INVENTION: Sequences and Antibodies Thereto

; NUMBER OF SEQUENCES: 31

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP

; STREET: Two Embarcadero Center, 8th Floor

; CITY: San Francisco

; STATE: CA

; COUNTRY: USA

; ZIP: 94111-3834

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/724,394A

; FILING DATE: 01-OCT-1996

; CLASSIFICATION: 536

; ATTORNEY/AGENT INFORMATION:

; NAME: Fitts, Renee A.

; REGISTRATION NUMBER: 35,136

; REFERENCE/DOCKET NUMBER: 017957-000100

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-576-0200

; TELEFAX: 415-576-0300

; INFORMATION FOR SEQ ID NO: 20:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 246240 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: not relevant

; TOPOLOGY: not relevant

; MOLECULE TYPE: cDNA

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: 1..246240

; OTHER INFORMATION: /note= "HLA-H. CONTIG"

US-08-724-394A-20

Query Match 7.6%; Score 229; DB 2; Length 246240;

Best Local Similarity 82.6%; Pred. No. 2e-52;

Matches 262; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

Qy 295 GTGATAACAGTGGCCATCAGCCGGTGCAGTGGCTAGTGGCTGAAATCCCAACACTTTGG 354

Db 2941 GTATTAAGGATGAACCCCGCTGGCGCGTGGCTCACCTGTAATCCAGCACTTTGG 2882

Qy 355 GAGCGGAGTGGTGGATCACTTGAAGTTCAGAGTTCGAGACCATCTCGCCAAACATGG 414

Db 2881 GAGCGGAGGCGGTAGATCACTTGAAGTTCAGAGTTCGAGACCATCTCGCCAAACATGG 2822

Qy 415 TGAACCTGTCTACTATAAAATATAAAATAGCTGGGCTAGTGGCTGCGTACCTGTA 474

Db 2821 CGAACCCTGTCTACTATAAAATATAAAATAGCTGGGCTAGTGGCTGCGGAGCACTGTA 2762

Qy 475 ATCCAGATATTCGGAGTTGAGCGAGGAGAAATCGCTTGAACCCGGGAGGCAATGTTG 534

Db 2761 ATCCAGCTGCTTGGGAGACTGAGCGAGGAGAAATCGTTTGAACCTGGGACGTTGG 2702

Qy 535 CAGTGAACCAAGACTGTGCCACTGCACCTCAGTCTGGGCAACAGAGTGAGCCTTCCATCTC 594

Db 2701 CAGTGAGCCAAAGACTGACCACTGCACCTCAGCTGGGCGACAGGGTGAGACAACTCTC 2642

Qy 595 AAACAACAAACAAAAA 611

Db 2641 AAAAAATAAAAAATAAA 2625

RESULT 15

US-08-724-394A-21/c

; Sequence 21, Application US/08724394A

; Patent No. 5872237

; GENERAL INFORMATION:

; APPLICANT: Feder, John N.

; APPLICANT: Kronmal, Gregory S.

; APPLICANT: Lauer, Peter M.

; APPLICANT: Ruddy, David A.

; APPLICANT: Thomas, Winston

; APPLICANT: Tsuchihashi, Zenta

; APPLICANT: Wolff, Roger K.

; TITLE OF INVENTION: Megabase Transcript Map: No. 5872237el

; TITLE OF INVENTION: Sequences and Antibodies Thereto

; NUMBER OF SEQUENCES: 31

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP

; STREET: Two Embarcadero Center, 8th Floor

; CITY: San Francisco

; STATE: CA

; COUNTRY: USA

; ZIP: 94111-3834

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/724,394A

; FILING DATE: 01-OCT-1996

; CLASSIFICATION: 536

; ATTORNEY/AGENT INFORMATION:

; NAME: Fitts, Renee A.

; REGISTRATION NUMBER: 35,136

; REFERENCE/DOCKET NUMBER: 017957-000100

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-576-0200

; TELEFAX: 415-576-0300

; INFORMATION FOR SEQ ID NO: 21:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 246240 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: not relevant

; TOPOLOGY: not relevant

; MOLECULE TYPE: cDNA

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: 1..246240

; OTHER INFORMATION: /note= "HLA-H. CONTIG"

US-08-724-394A-21

Query Match 7.6%; Score 229; DB 2; Length 246240;

Best Local Similarity 82.6%; Pred. No. 2e-52;

Matches 262; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

Qy 295 GTGATAACAGTGGCCATCAGCCGGTGCAGTGGCTAGTGGCTGAAATCCCAACACTTTGG 354

Db 2941 GTATTAAGGATGAACCCCGCTGGCGCGTGGCTCACCTGTAATCCAGCACTTTGG 2882

Qy 355 GAGCGGAGTGGTGGATCACTTGAAGTTCAGAGTTCGAGACCATCTCGCCAAACATGG 414

Db 2881 GAGCGGAGGCGGTAGATCACTTGAAGTTCAGAGTTCGAGACCATCTCGCCAAACATGG 2822

Qy 415 TGAACCTGTCTACTATAAAATATAAAATAGCTGGGCTAGTGGCTGCGTACCTGTA 474

Db 2821 CGAACCCTGTCTACTATAAAATATAAAATAGCTGGGCTAGTGGCTGCGGAGCACTGTA 2762

Qy 475 ATCCAGATATTCGGAGTTGAGCGAGGAGAAATCGCTTGAACCCGGGAGGCAATGTTG 534

Db 2761 ATCCAGCTGCTTGGGAGACTGAGCGAGGAGAAATCGTTTGAACCTGGGACGTTGG 2702

Qy	415	TGAAACCCCTGCTCTACTATAAAATATAAAATTTAGCTGGGCATGGTGGTACCTGTA	474
Db	2821	CGAAACCCCTGCTCTACTATAAAATATAAAATTTAGCTGGGCATGGTGGCAGCACCTGTA	2762
Qy	475	ATCCAGATACTTGGAGGTTGAGGCAGGAGAAATCGCTTGAAACCGGAGGCAGATGTTG	534
Db	2761	ATCCAGCTGCTTGGGAGACTGAGGCAGGAGAAATCGTTTGAACCTGGGACGTTGGGATTG	2702
Qy	535	CAGTGAACCAAGACTGTGCCACTGCCACTCCAGTCTGGGCAACAGAGTGAGCCTCCATCTC	594
Db	2701	CAGTGAGCCAGACTGCACCACTGCCACTCCAGCCTGGGCGACAGGGTGAGACAACTCTC	2642
Qy	595	AAACAAACAAACAAAA	611
Db	2641	AAAAATAAAAATATA	2625

Search completed: September 17, 2004, 19:04:56  
 Job time : 202.326 secs

***This Page Blank (uspto)***

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 17, 2004, 14:58:20 ; Search time 1330.64 Seconds  
(without alignments)  
11383.562 Million cell updates/sec

Title: US-10-669-693-3\_COPY\_1980\_4980

Perfect score: 3001

Sequence: 1 cttcttctctctaccaccc.....atcctggctaacacagtga 3001

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3327077 seqs, 2523723180 residues

Total number of hits satisfying chosen parameters: 6654154

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:\*

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*
- 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*
- 6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq.\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq.\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq.\*
- 11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq.\*
- 12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*
- 13: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*
- 14: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq.\*
- 15: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq.\*
- 16: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq.\*
- 17: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*
- 18: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*
- 19: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3001	100.0	17752	9	US-09-748-127-3
2	3001	100.0	17752	13	Sequence 3, Appli
C 3	282	9.4	61103	13	Sequence 58, Appl
C 4	253	8.4	16747	9	Sequence 3354, Ap
C 5	253	8.4	16747	16	Sequence 3354, Ap
C 6	248.4	8.3	122656	17	Sequence 846, App
C 7	248.4	8.3	137000	16	Sequence 11, Appl
C 8	247.6	8.2	786431	9	Sequence 1542, Ap
C 9	245.6	8.2	96594	15	Sequence 3, Appli
C 10	245.6	8.2	96594	12	Sequence 154, App
C 11	244.8	8.2	10500	15	Sequence 13, Appl
C 12	243	8.1	546	13	Sequence 256849
C 13	243	8.1	546	16	Sequence 256849
C 14	242.8	8.1	1243	13	Sequence 123618,

15	242.8	8.1	1243	13	US-10-027-632-123619	Sequence 123619,
16	242.8	8.1	1243	16	US-10-027-632-123618	Sequence 123618,
17	242.8	8.1	1243	16	US-10-027-632-123619	Sequence 123619,
C 18	242.6	8.1	2809	13	US-10-027-632-111984	Sequence 111984,
C 19	242.6	8.1	2809	13	US-10-027-632-111985	Sequence 111985,
C 20	242.6	8.1	2809	13	US-10-027-632-111986	Sequence 111986,
C 21	242.6	8.1	2809	16	US-10-027-632-111984	Sequence 111984,
C 22	242.6	8.1	2809	16	US-10-027-632-111985	Sequence 111985,
C 23	242.6	8.1	2809	16	US-10-027-632-111986	Sequence 111986,
24	242.4	8.1	31764	16	US-10-085-117-16	Sequence 16, Appl
25	241.8	8.1	162025	13	US-09-834-700-13	Sequence 13, Appl
26	241.8	8.1	162025	13	US-09-834-700-14	Sequence 14, Appl
27	241.8	8.1	162025	13	US-09-834-700-17	Sequence 17, Appl
28	241.8	8.1	162025	13	US-09-834-700-18	Sequence 18, Appl
29	241.8	8.1	162025	15	US-10-272-665-36	Sequence 36, Appl
30	241.8	8.1	162025	15	US-10-272-665-35	Sequence 35, Appl
31	241.8	8.1	162025	15	US-10-273-321-35	Sequence 35, Appl
32	241.8	8.1	162025	15	US-10-273-321-36	Sequence 36, Appl
33	241.8	8.1	162025	15	US-10-272-756-35	Sequence 35, Appl
34	241.8	8.1	162025	15	US-10-272-756-36	Sequence 36, Appl
35	241.8	8.1	162025	16	US-10-273-228-35	Sequence 35, Appl
36	241.8	8.1	162025	16	US-10-273-228-36	Sequence 36, Appl
C 37	241	8.0	160482	13	US-10-087-192-100	Sequence 100, App
C 38	240.4	8.0	95960	13	US-10-087-192-1384	Sequence 1384, Ap
C 39	240.4	8.0	95960	13	US-10-087-192-1390	Sequence 1390, Ap
C 40	240.2	8.0	98829	15	US-10-017-724-3	Sequence 3, Appli
C 41	240	8.0	154902	17	US-10-450-826-105	Sequence 105, App
C 42	238.8	8.0	9297	12	US-10-160-388-1	Sequence 1, Appli
C 43	238.8	8.0	10953	9	US-09-764-846-313	Sequence 313, App
C 44	238.8	8.0	10953	15	US-10-091-483-313	Sequence 313, App
C 45	238.8	8.0	22645	10	US-09-764-891-7673	Sequence 7673, Ap

ALIGNMENTS

RESULT 1  
US-09-748-127-3  
; Sequence 3, Application US/09748127  
; Patent No. US20020076774A1  
; GENERAL INFORMATION:  
; APPLICANT: YAN Chunhua et al  
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING  
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN  
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS,  
; TITLE OF INVENTION: AND USES THEREOF  
; FILE REFERENCE: CL000685  
; CURRENT APPLICATION NUMBER: US/09/748,127  
; CURRENT FILING DATE: 2001-06-11  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 17752  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(17752)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-748-127-3

Query Match 100.0%; Score 3001; DB 9; Length 17752;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 3001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CTTTCTCTCTCTACCCACGCTGAGTACGACCGGTGTACCATCTACCT 60  
DB 1980 CTTTCTCTCTCTCTACCCACGCTGAGTACGACCGGTGTACCATCTACCT 2039  
QY 61 GGGACCTCGCGGCTGTGGTGTCTCTGTTGGCAGGAGGCTGTGGGAGCCCTGGG 120  
DB 2040 GGGACCTCGCGGCTGTGGTGTCTCTGTTGGCAGGAGGCTGTGGGAGCCCTGGG 2099

QY	121	AGGTCAGGCTGAGGAGTTACGGCGCCGGGAAACCGTAGGGATGCTGGGAAGGAGCTTTTGA	180	Db	3180	TGACCTTCAGAGAAGCAATTCATCCATTCAACTGATGAATTTTCAGACTGGGACCGCTGGC	3239
Db	2100	AGGTGAGGCTGAGGAGTTACGGCGCCGGGAAACCGTAGGGATGCTGGGAAGGAGCTTTTGA	2159	QY	1261	TCATGCCCTGTAAATCCACGACATTTTGGAAAGGCTGAATGGGAGGATGACTTGTAGCCCTAGGC	1320
QY	181	TGGCCATGTTAGTCAAGGGCTGTAGGCCCTCCGCTACAGCCTGCCACCACTTACTTGG	240	Db	3240	TCATGCCCTGTAAATCCACGACATTTTGGAAAGGCTGAATGGGAGGATGACTTGTAGCCCTAGGC	3299
Db	2160	TGGCCATGTTAGTCAAGGGCTGTAGGCCCTCCGCTACAGCCTGCCACCACTTACTTGG	2219	QY	1321	ATTGTGTACAGCCTTGGGCAACATGCTGAGACCTTGCCTCCACAAAAAACAACAACAAC	1380
QY	241	TGTGTGACCTTTGACATGGCTTAGTCCCTCTGTGTGCTCATCTGTCAAATGGAGTGATA	300	Db	3300	ATTGTGTGACAGCCTTGGGCAACATGCTGAGACCTTGCCTCCACAAAAAACAACAACAAC	3359
Db	2220	TGTGTGACCTTTGACATGGCTTAGTCCCTCTGTGTGCTCATCTGTCAAATGGAGTGATA	2279	QY	1381	AAAAAATCAATTAATCTTGTTACCATGGGTACCAAGGTACATAGAAATGACTCAGGCAGATA	1440
QY	301	ACAGTGCCCATCAGCCGGGTGCAAGTGGCTAGTGGCTGTAGTCCCTGAAATCCCAACACATTTGGGAGGCG	360	Db	3360	AAAAAATCAATTAATCTTGTTACCATGGGTACCAAGGTACATAGAAATGACTCAGGCAGATA	3419
Db	2280	ACAGTGCCCATCAGCCGGGTGCAAGTGGCTAGTGGCTGTAGTCCCTGAAATCCCAACACATTTGGGAGGCG	2339	QY	1441	TGGTGTCTCTCTACTGTGGAGAGCGGGCTTACTTGCAGTAAGACAATAGAGGGAG	1500
QY	361	GAGTGGGTGGATCACTTCAGCTCAGGAGTTGAGACACAGCCTGGCCAAACATGGTGAAC	420	Db	3420	TGGTGTCTCTCTACTGTGGAGAGCGGGCTTACTTGCAGTAAGACAATAGAGGGAG	3479
Db	2340	GAGTGGGTGGATCACTTCAGCTCAGGAGTTGAGACACAGCCTGGCCAAACATGGTGAAC	2399	QY	1501	GGAAATATAATCTTAAATCGAGGTACAGATTTGAGAGCAAAACACAGGGCACAGGCATAT	1560
QY	421	CCTGTCTCTACTAAAAATATAAAATTTAGCTGGGCAATGGTGGCTGATCCTGTAATCCCA	480	Db	3480	GGAAATATAATCTTAAATCGAGGTACAGATTTGAGAGCAAAACACAGGGCACAGGCATAT	3539
Db	2400	CCTGTCTCTACTAAAAATATAAAATTTAGCTGGGCAATGGTGGCTGATCCTGTAATCCCA	2459	QY	1561	GTACGAGGGTAAAGAGGGAATCAGGGAAGCTTCTCAGAGAAGGTGACATTTAAGCCGGG	1620
QY	481	GATACTTGGAGGTTGAGGACGAGAAATGCTTTGAACCCGGGAGGAGATGTTGCAAGTGA	540	Db	3540	GTACGAGGGTAAAGAGGGAATCAGGGAAGCTTCTCAGAGAAGGTGACATTTAAGCCGGG	3599
Db	2460	GATACTTGGAGGTTGAGGACGAGAAATGCTTTGAACCCGGGAGGAGATGTTGCAAGTGA	2519	QY	1621	ACATGAAGGATGAACGAGTTAGTTTACCAAGGATGGATGGAAAGGGGTGAGAGTGATGG	1680
QY	541	ACCAAGACTGTGCCACTGCATCTCAGTCTGGGCAACAGAGTGAGCTTCCATCTCAAACAA	600	Db	3600	ACATGAAGGATGAACGAGTTAGTTTACCAAGGATGGATGGAAAGGGGTGAGAGTGATGG	3659
Db	2520	ACCAAGACTGTGCCACTGCATCTCAGTCTGGGCAACAGAGTGAGCTTCCATCTCAAACAA	2579	QY	1681	AGGCACAGGGAACCTGCAGGATCATAGCCCTAGACAGGGATCCTGCAGCCCTTTGAGGAAG	1740
QY	601	ACAAAACAAAAGCAGTGCCTCATCATGTAGGATTTGAGTGAATGAGTGAGGACTCAGCCTTG	660	Db	3660	AGGCACAGGGAACCTGCAGGATCATAGCCCTAGACAGGGATCCTGCAGCCCTTTGAGGAAG	3719
Db	2580	ACAAAACAAAAGCAGTGCCTCATCATGTAGGATTTGAGTGAATGAGTGAGGACTCAGCCTTG	2639	QY	1741	TGAGAAGAAGCAGCCAGTCGTAGTGGGTAAAGTAAACAAAGCTGAGAAAGCCAGGGAAT	1800
QY	661	TGCAAAAGTGAAGTCACTAATCACAGGTTGTAGTATCAGTATACCATCAATCATATCC	720	Db	3720	TGAGAAGAAGCAGCCAGTCGTAGTGGGTAAAGTAAACAAAGCTGAGAAAGCCAGGGAAT	3779
Db	2640	TGCAAAAGTGAAGTCACTAATCACAGGTTGTAGTATCAGTATACCATCAATCATATCC	2699	QY	1801	CCCTGGTCACTGACAGGCTGTGAGTCAAGTCAAGTGTGTTGGCTTTTGTCTTCTGGG	1860
QY	721	AGTTAAAGCCCTCAGGGTTTCAGAAAGATGCCGAGCGCTTTCAGAGTGTCTGGGGATTGGT	780	Db	3780	CCCTGGTCACTGACAGGCTGTGAGTCAAGTCAAGTGTGTTGGCTTTTGTCTTCTGGG	3839
Db	2700	AGTTAAAGCCCTCAGGGTTTCAGAAAGATGCCGAGCGCTTTCAGAGTGTCTGGGGATTGGT	2759	QY	1861	AGCAGTCTGATTTTAAAGCAGGGAACAGCTGTATTTTACAGATTTGGGAAGATCCTGTGTTGCT	1920
QY	781	GGGCAAGCCCTCGAATTAATAGAAACAGTTCTCTGTATTACACAGAAAGCAGAGGCCCA	840	Db	3840	AGCAGTCTGATTTTAAAGCAGGGAACAGCTGTATTTTACAGATTTGGGAAGATCCTGTGTTGCT	3899
Db	2760	GGGCAAGCCCTCGAATTAATAGAAACAGTTCTCTGTATTACACAGAAAGCAGAGGCCCA	2819	QY	1921	GCCTGAAGGGGATGAAAATCGAGGCTAGGAGCCAGGTCATAGGGAGGATCCAGGGTGA	1980
QY	841	TGCTGGGTGTGCCAGGAACTCAGTAGTAATAAGACAGCAGCACCGGTCTGCTTCCCAGC	900	Db	3900	GCCTGAAGGGGATGAAAATCGAGGCTAGGAGCCAGGTCATAGGGAGGATCCAGGGTGA	3959
Db	2820	TGCTGGGTGTGCCAGGAACTCAGTAGTAATAAGACAGCAGCACCGGTCTGCTTCCCAGC	2879	QY	1981	TGGGGAGGCTGGGAGGTCCCGGTGTATGGACACAGGCTGGGGCCAGGGATGGGAGGAA	2040
QY	901	GCACCTAGGCCAGTGGGGAACAGACTCACACACAGTCCAGCCACAGAGTGTGTCAGGGC	960	Db	3960	TGGGGAGGCTGGGAGGTCCCGGTGTATGGACACAGGCTGGGGCCAGGGATGGGAGGAA	4019
Db	2880	GCACCTAGGCCAGTGGGGAACAGACTCACACACAGTCCAGCCACAGAGTGTGTCAGGGC	2939	QY	2041	GGAGTAATTCGGGAGAGCCCTCTGGCCCTGCGCGGTGATGGACACAGGCTGGGGCCAGGGAT	2100
QY	961	CAAGATGGGGAAGCACCGGGAGAAAGTTCAGGGTGGGATGGGAGGGGTGAGGGCAAGAG	1020	Db	4020	GGAGTAATTCGGGAGAGCCCTCTGGCCCTGCGCGGTGATGGACACAGGCTGGGGCCAGGGAT	4079
Db	2940	CAAGATGGGGAAGCACCGGGAGAAAGTTCAGGGTGGGATGGGAGGGGTGAGGGCAAGAG	2999	QY	2101	AGAGGAGAGATGCTTAGGCCACTTTTGGAAACACAGTAGGGCCAGGACACAGGAGACACCCAG	2160
QY	1021	GGGTGAGGCCAGCTGAGGGAAGCCCTGGGACTGTAGGATTTAGAGGAGGTACCTGAC	1080	Db	4080	AGAGGAGAGATGCTTAGGCCACTTTTGGAAACACAGTAGGGCCAGGACACAGGAGACACCCAG	4139
Db	3000	GGGTGAGGCCAGCTGAGGGAAGCCCTGGGACTGTAGGATTTAGAGGAGGTACCTGAC	3059	QY	2161	GGGAGGTGCCCAAGAGACCCAGCAGGCTGTGATTTGGACAGGGAAGGTCTGTCTGGAGCA	2220
QY	1081	CCGGCATGTTTGGTGAGGGAGATTGAGGAAGTCTTCTCTGGGAAGAGAGGCTGTCCGAGCTG	1140	Db	4140	GGGAGGTGCCCAAGAGACCCAGCAGGCTGTGATTTGGACAGGGAAGGTCTGTCTGGAGCA	4199
Db	3060	CCGGCATGTTTGGTGAGGGAGATTGAGGAAGTCTTCTCTGGGAAGAGAGGCTGTCCGAGCTG	3119	QY	2221	GGTGTCTTGGATTAAGGGAGGAAAATGGTGCAGTTTCCATCTCTCCCTCTCTGTGTCAACC	2280
QY	1141	AGACTCAATAAGATGAGTGGGAGGGTGTTCAGGCAGAAAGACACAGCACCTTACAAAGCA	1200	Db	4200	GGTGTCTTGGATTAAGGGAGGAAAATGGTGCAGTTTCCATCTCTCTCTCTCTGTGTCAACC	4259
Db	3120	AGACTCAATAAGATGAGTGGGAGGGTGTTCAGGCAGAAAGACACAGCACCTTACAAAGCA	3179	QY	2281	TCTAAAATCAATGGGGCACAGGACCCAGTGGGACTCCATTAATGATGGGATGGGTGGATG	2340
QY	1201	TGACTTTGAGAGAGCAATTCATCCATTCAACTGATGAATTTTTCAGACTGGGCAAGCTGGC	1260				



```
Db 4260 TCTAACTACATGGGGCACAGGACCCAGTGGGACTCCCATAAATGATGGGTGGATG 4319
QY 2341 GAAGGAAGGAGGAGGAACAACTCTTCAATCATCTCTGTTATTTACAGAACAGGCCAGG 2400
Db 4320 GAAGGAAGGAGGAGGAACAACTCTTCAATCATCTCTGTTATTTACAGAACAGGCCAGG 4379
QY 2401 TGGCGTGTCTACCGTTGCGCATTTAGCACATTTGGGAGGCTGAGGTGGGTGGATTACCTCA 2460
Db 4380 TGGCGTGTCTACCGTTGCGCATTTAGCACATTTGGGAGGCTGAGGTGGGTGGATTACCTCA 4439
QY 2461 GGTCAAGGAGTTCAAGACAGCGCTAGACAAAGCTAGAGAAACCCCATCTCTACTGAAGATAT 2520
Db 4440 GGTCAAGGAGTTCAAGACAGCGCTAGACAAAGCTAGAGAAACCCCATCTCTACTGAAGATAT 4499
QY 2521 AAAATTAGCTGGGCGTAGTGCATATGCTGTATCCAGCTAGTCCGGAAGCTGAGGCA 2580
Db 4500 AAAATTAGCTGGGCGTAGTGCATATGCTGTATCCAGCTAGTCCGGAAGCTGAGGCA 4559
QY 2581 GGAGAAATCGCTTGAACCCGAGAGGCGAGAGGTTGGGTGAGCTGAGATCGTGCCATTGCGAC 2640
Db 4560 GGAGAAATCGCTTGAACCCGAGAGGCGAGAGGTTGGGTGAGCTGAGATCGTGCCATTGCGAC 4619
QY 2641 TCCAGCTGGGTGACAAAGCAAGACCTCGTCTCAATAATAATAATAATAATAATAATAATA 2700
Db 4620 TCCAGCTGGGTGACAAAGCAAGACCTCGTCTCAATAATAATAATAATAATAATAATAATA 4679
QY 2701 GGAGCTGGGTGATCCAGCTACCTACTTTTCAGAGAAATGCTACTCCCTTACCCAAAGGC 2760
Db 4680 GGAGCTGGGTGATCCAGCTACCTACTTTTCAGAGAAATGCTACTCCCTTACCCAAAGGC 4739
QY 2761 AAAGATGGGAGAACCAAGTTGATATGCAATTTATGAGCACTTACTGAGTCCCTCATCCC 2820
Db 4740 AAAGATGGGAGAACCAAGTTGATATGCAATTTATGAGCACTTACTGAGTCCCTCATCCC 4799
QY 2821 TGGGCTAGGCTGGAATGGATCAGATCGAGCTGAAGAGTCCCGCTCAGGGAACCTCACT 2880
Db 4800 TGGGCTAGGCTGGAATGGATCAGATCGAGCTGAAGAGTCCCGCTCAGGGAACCTCACT 4859
QY 2881 AGAAGAGGAGGAATCGCGGGCGGTGGCTCAGCGCTGTAATCCCAACATTTGGG 2940
Db 4860 AGAAGAGGAGGAATCGCGGGCGGTGGCTCAGCGCTGTAATCCCAACATTTGGG 4919
QY 2941 AGGCTAGGTTGGGTGGATCAAGGTCAGGAGATCGAGACCATCTCGCTAAACAGTGA 3000
Db 4920 AGGCTAGGTTGGGTGGATCAAGGTCAGGAGATCGAGACCATCTCGCTAAACAGTGA 4979
QY 3001 A 3001
Db 4980 A 4980
```

RESULT 2  
US-10-669-693-3

```
; Sequence 3, Application US/10669693
; Publication No. US2004004313A1
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS,
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL000685CON
; CURRENT APPLICATION NUMBER: US/10/669,693
; PRIOR FILING DATE: 2003-09-25
; PRIOR FILING DATE: 2000-12-27
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 17752
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
```

```
; NAME/KEY: misc_feature
; LOCATION: (1)...(17752)
; OTHER INFORMATION: n = A,T,C or G
US-10-669-693-3
```

```
Query Match 100.0%; Score 3001; DB 13; Length 17752;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTTCTCTCTCCCTACCCAGCTGAGTAAGAGTACGACCGGTGTTACCATCTACT 60
Db 1980 CTTTCTCTCTCCCTACCCAGCTGAGTAAGAGTACGACCGGTGTTACCATCTACT 2039
QY 61 GGGACCTTGGCGGCTGTGTGTCCTGTTGGGACGAGGCTGTGGGGAGGCCCTGGG 120
Db 2040 GGGACCTTGGCGGCTGTGTGTCCTGTTGGGACGAGGCTGTGGGGAGGCCCTGGG 2099
QY 121 AGGTCAGGCTGAGGAGTTTCAGCGGCGGGGAAACCGTAGCGATGCTGGGAAGGACCTTTTGA 180
Db 2100 AGGTCAGGCTGAGGAGTTTCAGCGGCGGGGAAACCGTAGCGATGCTGGGAAGGACCTTTTGA 2159
QY 181 TGGCCATGTTAAGTCAAGGGCTGTAGGCCCTCCGCTCACAGCCTGCCACCTTACTGG 240
Db 2160 TGGCCATGTTAAGTCAAGGGCTGTAGGCCCTCCGCTCACAGCCTGCCACCTTACTGG 2219
QY 241 TGTGTGACCTTTGACATGAGCTTAGTCCCTCTGTGTGCTCATCTGTCAAATGGAGTGATA 300
Db 2220 TGTGTGACCTTTGACATGAGCTTAGTCCCTCTGTGTGCTCATCTGTCAAATGGAGTGATA 2279
QY 301 ACAGTGGCCATCAGCGGGTGCAGTGGCTAGTGGCTGAAATCCCAACATTTGGGAGGCG 360
Db 2280 ACAGTGGCCATCAGCGGGTGCAGTGGCTAGTGGCTGAAATCCCAACATTTGGGAGGCG 2339
QY 361 GAGTGGGTGGATCACATTGAGGTCAGGAGTTCAGAGCCAGCTGGCCAAACATGTTGAAC 420
Db 2340 GAGTGGGTGGATCACATTGAGGTCAGGAGTTCAGAGCCAGCTGGCCAAACATGTTGAAC 2399
QY 421 CTTGCTCTACTAAATAATAATAAATTAAGTGGGCTAGTGGTGGCTGTAATCCCA 480
Db 2400 CTTGCTCTACTAAATAATAATAAATTAAGTGGGCTAGTGGTGGCTGTAATCCCA 2459
QY 481 GATACCTGGGAGGTTGAGGAGGAGGAGTTCGCTTGAACCCGGGAGGAGATGTTGAGTGA 540
Db 2460 GATACCTGGGAGGTTGAGGAGGAGGAGTTCGCTTGAACCCGGGAGGAGATGTTGAGTGA 2519
QY 541 ACCAAGACTGTGCCACTGCACTCCAGTCTGGGCAACAGAGTGAGCTCCATCTCAAACAA 600
Db 2520 ACCAAGACTGTGCCACTGCACTCCAGTCTGGGCAACAGAGTGAGCTCCATCTCAAACAA 2579
QY 601 ACAAACAAAGCAGTGGCCCATCATGTAGATTGAGTTGAGTGGAGGACTGAGCCTTG 660
Db 2580 ACAAACAAAGCAGTGGCCCATCATGTAGATTGAGTTGAGTGGAGGACTGAGCCTTG 2639
QY 661 TGCAAGTGGAGCTCACTAATCAGCAGGTTGTAGTATCATGTATACCATCAATGATCC 720
Db 2640 TGCAAGTGGAGCTCACTAATCAGCAGGTTGTAGTATCATGTATACCATCAATGATCC 2699
QY 721 AGGTAAGCCCTCAGGAGTTTCAGAAAGATGCGGAGCGCTTTCAAGGTGCTGGGATTGGT 780
Db 2700 AGGTAAGCCCTCAGGAGTTTCAGAAAGATGCGGAGCGCTTTCAAGGTGCTGGGATTGGT 2759
QY 781 GGGCAAGCCCTCGAATAATAGAAACAGTTCTCTGTATTACAAACAGAAAGCAGGAGGCCCA 840
Db 2760 GGGCAAGCCCTCGAATAATAGAAACAGTTCTCTGTATTACAAACAGAAAGCAGGAGGCCCA 2819
QY 841 TGCTGGGTGCTGCCAGGAACTCAGTAGTAATTAAGACAGCAGCGGTGCTGCTCCCGAGC 900
Db 2820 TGCTGGGTGCTGCCAGGAACTCAGTAGTAATTAAGACAGCAGCGGTGCTGCTCCCGAGC 2879
QY 901 GCACCTAGGCGAGTGGGGAACAGAGCTCACCAACAGTCCCGCCAGAGAGTGCTCAGGCG 960
Db 2880 GCACCTAGGCGAGTGGGGAACAGAGCTCACCAACAGTCCCGCCAGAGAGTGCTCAGGCG 2939
```

QY 961 CAAGATGGGGAAGCA CGGGAGAAAGGT CAGGTTGGATGGGAGGGGT CAGGGCAAG 1020  
 DB 2940 CAAGATGGGGAAGCA CGGGAGAAAGGT CAGGTTGGATGGGAGGGGT CAGGGCAAG 2999  
 QY 1021 GGGTCAGGGCCAGGCTGAGGGAAGCCCTGGGACTGTAGGAAATTTAGAGGAGTACTGTAC 1080  
 DB 3000 GGGTCAGGGCCAGGCTGAGGGAAGCCCTGGGACTGTAGGAAATTTAGAGGAGTACTGTAC 3059  
 QY 1081 CCGGCATGTTTGGTGGAGGAGATTCAGGAAGTCTTCTTGGAAAGAGAGGCTGTCGAGCTG 1140  
 DB 3060 CCGGCATGTTTGGTGGAGGAGATTCAGGAAGTCTTCTTGGAAAGAGAGGCTGTCGAGCTG 3119  
 QY 1141 AGACTCATAGATGAGTGGGAGGAGTTCAGGCAAGAGACCCAGCACCTACAAAGCA 1200  
 DB 3120 AGACTCATAGATGAGTGGGAGGAGTTCAGGCAAGAGACCCAGCACCTACAAAGCA 3179  
 QY 1201 TGACTTTGAGAGAAGCATTCATCCATTCACCTGATGAAATTTTCAGACTGGGCAAGCTGGC 1260  
 DB 3180 TGACTTTGAGAGAAGCATTCATCCATTCACCTGATGAAATTTTCAGACTGGGCAAGCTGGC 3239  
 QY 1261 TCATGCCCTGTAATCCAGCACTTTGGAAAGCTGGAATGGGAGAGTACTGTAGGCTTAGGC 1320  
 DB 3240 TCATGCCCTGTAATCCAGCACTTTGGAAAGCTGGAATGGGAGAGTACTGTAGGCTTAGGC 3299  
 QY 1321 ATTTGTGACAAGCTGGGCAACATGTTGAGACCTCTCCACAAACAAACAAACAAAC 1380  
 DB 3300 ATTTGTGACAAGCTGGGCAACATGTTGAGACCTCTCCACAAACAAACAAACAAAC 3359  
 QY 1381 AAAAAATCATTAATCTGTTGATCCATGGGTACAGGTACATAGAAATGACTCAGGCAGATA 1440  
 DB 3360 AAAAAATCATTAATCTGTTGATCCATGGGTACAGGTACATAGAAATGACTCAGGCAGATA 3419  
 QY 1441 TGGTGTCTCTCTACTGTGGAGAGCGGGCTTATACTGAGTAAGACAAATAGAGGAG 1500  
 DB 3420 TGGTGTCTCTCTACTGTGGAGAGCGGGCTTATACTGAGTAAGACAAATAGAGGAG 3479  
 QY 1501 GGAATATAATCTTAAATGAGGTACAGTTTGGAGCAACACAGGGCACAGGCATAT 1560  
 DB 3480 GGAATATAATCTTAAATGAGGTACAGTTTGGAGCAACACAGGGCACAGGCATAT 3539  
 QY 1561 GTACGAGGTAAAGAGGGAATCAGGGAAGGCTTCTCAGAGAAGGTGACATTTAAGCCGG 1620  
 DB 3540 GTACGAGGTAAAGAGGGAATCAGGGAAGGCTTCTCAGAGAAGGTGACATTTAAGCCGG 3599  
 QY 1621 ACATGAAGGATGAACAGTTAGTTTACCAAGGATGGGATGGAAAGGGGTGAGAGTATG 1680  
 DB 3600 ACATGAAGGATGAACAGTTAGTTTACCAAGGATGGGATGGAAAGGGGTGAGAGTATG 3659  
 QY 1681 AGGCAGGGAACTGCAGGATCATAGGCTTAGACAGGGATCTGACGCCCTTTAGGAAG 1740  
 DB 3660 AGGCAGGGAACTGCAGGATCATAGGCTTAGACAGGGATCTGACGCCCTTTAGGAAG 3719  
 QY 1741 TGAGAGAAGACAGCGCAGTCTGAGTGGGTAAAGTAAACAAAGCTGAGAAACCCAGGGAAT 1800  
 DB 3720 TGAGAGAAGACAGCGCAGTCTGAGTGGGTAAAGTAAACAAAGCTGAGAAACCCAGGGAAT 3779  
 QY 1801 CCTGTGTATGCAAGGCTGTGAGTCAAGTCAAGTGTGTTGGGCTTTGTTTCTCTGG 1860  
 DB 3780 CCTGTGTATGCAAGGCTGTGAGTCAAGTCAAGTGTGTTGGGCTTTGTTTCTCTGG 3839  
 QY 1861 AGCAGTCGATTTTAAAGCAGGGAACAGCTGTATTCAGAGTTGGGAAGATCTGTGTTGCT 1920  
 DB 3840 AGCAGTCGATTTTAAAGCAGGGAACAGCTGTATTCAGAGTTGGGAAGATCTGTGTTGCT 3899  
 QY 1921 GCCTGAAGGGATGAATCTGAGGCTAGGAGCCAGGGTATAGGGAGGATCCAGGCTGA 1980  
 DB 3900 GCCTGAAGGGATGAATCTGAGGCTAGGAGCCAGGGTATAGGGAGGATCCAGGCTGA 3959  
 QY 1981 TGGGGAGGCTGGAGGTCGCGGTGATGGACACAGGGCTGGGGCAGGGGATGGGGAGAA 2040  
 DB 3960 TGGGGAGGCTGGAGGTCGCGGTGATGGACACAGGGCTGGGGCAGGGGATGGGGAGAA 4019  
 QY 2041 GGAGTAATGGGAGAGGCTGGGGCTCTGGCCGAGGAATGATGTTGGGCTGAACACAGG 2100

DB 4020 GGAGTAATTTGGAGAGGCTTGGGGCTCTGGCCAGGAATGGATGGTGGCTGAAACAGGG 4079  
 QY 2101 AGAGAGAGATGCTTTAGGCCACTTTGGGAACACAGTAGGCAAGACAGGAGACACCCAG 2160  
 DB 4080 AGAGAGAGATGCTTTAGGCCACTTTGGGAACACAGTAGGCAAGACAGGAGACACCCAG 4139  
 QY 2161 GGAAGTCCCAAGAGACACAGCAGGCTGGCATTGGACAGGAAAGGTCTGTCTCGAGCA 2220  
 DB 4140 GGAAGTCCCAAGAGACACAGCAGGCTGGCATTGGACAGGAAAGGTCTGTCTCGAGCA 4199  
 QY 2221 GGTGTCTTGGATAAGGGAGGAAATGGTGCAGTTCATCTCTCTCTCTCTCTCTCTCTCT 2280  
 DB 4200 GGTGTCTTGGATAAGGGAGGAAATGGTGCAGTTCATCTCTCTCTCTCTCTCTCTCTCT 4259  
 QY 2281 TCTAAATCATGATGGGCAAGACCCAGTGGGACTCCATAAATGATGGATGGGTGGATG 2340  
 DB 4260 TCTAAATCATGATGGGCAAGACCCAGTGGGACTCCATAAATGATGGATGGGTGGATG 4319  
 QY 2341 GAAGGAAGGAAGGAGAAACAACTCTTCAATTCATCTGCTTATTTACAGAACAGGCCAGG 2400  
 DB 4320 GAAGGAAGGAAGGAGAAACAACTCTTCAATTCATCTGCTTATTTACAGAACAGGCCAGG 4379  
 QY 2401 TCGGGTGTCTACGCTTGGCCATTCTAGCACTTTGGAGGCTGAGGTGGGTGGATTTACTCA 2460  
 DB 4380 TCGGGTGTCTACGCTTGGCCATTCTAGCACTTTGGAGGCTGAGGTGGGTGGATTTACTCA 4439  
 QY 2461 GGTAGGAGTTCAAGACGAGCTAGACACCTGAGAAACCCCTCTCTACTGAGCATAT 2520  
 DB 4440 GGTAGGAGTTCAAGACGAGCTAGACACCTGAGAAACCCCTCTCTACTGAGCATAT 4499  
 QY 2521 AAAATTTAGCTGGGCTGAGTGGCATATGCTCTGTAATCCAGCTAGTGGGAAGCTGAGGCA 2580  
 DB 4500 AAAATTTAGCTGGGCTGAGTGGCATATGCTCTGTAATCCAGCTAGTGGGAAGCTGAGGCA 4559  
 QY 2581 GGAGAAATCGCTTTAAACCCAGAGGAGAGGTTGCGGTGAGCTGAGATCGTGCCATTTGCAC 2640  
 DB 4560 GGAGAAATCGCTTTAAACCCAGAGGAGAGGTTGCGGTGAGCTGAGATCGTGCCATTTGCAC 4619  
 QY 2641 TCCAGCTGGGTGACAAAGCAAGACCTCGCTCTCAATAATAATAATAATAATAATAACAGAA 2700  
 DB 4620 TCCAGCTGGGTGACAAAGCAAGACCTCGCTCTCAATAATAATAATAATAATAATAACAGAA 4679  
 QY 2701 GGAGCTGGGTCAATCCAGCTACTCTTTTTCAGGAGAAATGTACTCCCTTACCCAAAGGC 2760  
 DB 4680 GGAGCTGGGTCAATCCAGCTACTCTTTTTCAGGAGAAATGTACTCCCTTACCCAAAGGC 4739  
 QY 2761 AAAGGATGGAGAAACAGTTTGAATATGCAATTTATTTGAGCACTCTGAGTCTCTATCCC 2820  
 DB 4740 AAAGGATGGAGAAACAGTTTGAATATGCAATTTATTTGAGCACTCTGAGTCTCTATCCC 4799  
 QY 2821 TGGGCTAGGCTGGAATGGAATGAGCTCAGATGGAGCTGAGAGTCCCTTCAGGGAACCTCACT 2880  
 DB 4800 TGGGCTAGGCTGGAATGGAATGAGCTCAGATGGAGCTGAGAGTCCCTTCAGGGAACCTCACT 4859  
 QY 2881 AGAAAGAGAGAAATCGCGCGCGGTGGCTCACGCTGTAAATCCCAACACTTTTGGG 2940  
 DB 4860 AGAAAGAGAGAAATCGCGCGCGGTGGCTCACGCTGTAAATCCCAACACTTTTGGG 4919  
 QY 2941 AGGCTAGGTTGGTGGATCAAGAGTCAAGAGTCAAGAGATCGAGACCATCTCTGGCTTAAACAGTGA 3000  
 DB 4920 AGGCTAGGTTGGTGGATCAAGAGTCAAGAGTCAAGAGATCGAGACCATCTCTGGCTTAAACAGTGA 4979  
 QY 3001 A 3001  
 DB 4980 A 4980

RESULT 3  
 US-10-087-192-58/c  
 ; Sequence 58, Application US/10087192  
 ; Publication No. US20020182586A1  
 ; GENERAL INFORMATION:



[illegible]

```

; NAME/KEY: SITE
; LOCATION: (14067)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (14068)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (14069)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (14070)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (14071)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (14072)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (14073)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (14074)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (14075)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (14076)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (14077)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (14078)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (14079)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (14080)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (14081)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (14082)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (14083)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (14084)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (14085)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (14086)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (14087)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (14088)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (14089)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (14090)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE

```

```

; LOCATION: (14091)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (14092)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (14093)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (14094)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (14095)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (14096)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (14097)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE

Query Match      8.4%; Score 253; DB 9; Length 16747;
Best Local Similarity 27.2%; Pred. No. 1.9e-65;
Matches 659; Conservative 0; Mismatches 1740; Indels 22; Gaps 6;

Qy 301 ACAGTGGCCATCAGCGGGTGCAGTGGCTGAAATCCCAACACTTTGGGAGGCG 360
Db |||||
16152 AAAATGAGTTTAGGCTGGGCGGCTGCTAGGCTATATCCCAACACTTTGGGAGGCC 16093
Qy 361 GAGGTGGTGGATCATTGAGGTGAGAGTTGAGAGCAGCCTGGCCAAACATGTTGAAAC 420
Db |||||
16092 GAGGTGGTGGATCATTGAGGTGAGAGTTGAGAGCAGCCTGGCCAAACATGTTGAAAC 16033
Qy 421 CCTGCTCTACTAAATA-TAAAAATTAGTGGGCATGTTGGTGCCTACCTGTAATCCC 479
Db |||||
16032 TCTGCTCTACTAAATTACAAAAATTAGTGGGCATGTTGGCGGTGCTGTAATCCC 15973
Qy 480 AGATACTTGGGAGGTGAGGCGAGGAGATCGCTTGAACCCGGGAGGAGATGTTGCAGTG 539
Db |||||
15972 AGCTACTTGGGAGGCTGAGGCGAGAGATTCCTGTAACCCAGAGATGGAGGTTGTAGTG 15913
Qy 540 AACCAAGACTGTGCCACTGCACCTCCAGTCTGGGCAACA-GAGTGAGGCTCCATCTCAAAAC 598
Db |||||
15912 AGCCAAGACTGCGCCACTGCACCTAGCCTGGGCAACAACGAGCGGAACCTCGCGCTCAAAA 15853
Qy 599 AAACAAACAAAAGACAGTGCCTCATGTAGGATTGAGTGTGATTGAGTACGAGACTGAGCCT 658
Db |||||
15852 AAAAAAATAAATTGGAGGAGATGAGATTGTAATGTAATATATAGACATCATTTTA 15793
Qy 659 TGTGCAAGTGCAGCACTCACTAATCAACAGGTTGTAGTATCAGTGATAACCATCAATGAT 718
Db |||||
15792 TAGGAATAGTGAGTCTCTGAAGACCTCTATGATGGGTTAATTGGGGAAGATGTTTCATATT 15733
Qy 719 CCAGGTAAGCCCTGAGGTTTCAGAAAGATGCCGAGCGCTTTCAAGGTGCTGGGGATTG 778
Db |||||
15732 TCTGAGTCAGAAAAAATTTTACAGACAAATTTGCTGTCTCAGTTCAGAGAGTGGAGAACAGA 15673
Qy 779 GTGGCAAGCCCTCGAATAATAGAAACAGTTCTCTATTATACACAGAAAGAGGAGGCC 838
Db |||||
15672 ATGAGGCGAAGGGGATTGAGAGGTGTGCAATTATTTTCAGGTGAGAGGTAAGACAG 15613
Qy 839 CATGCTGGGTGCTGCCAGGAACCTCAGTAGTAACTAAGACAGACACCGGTGCTGCTCCCA 898
Db |||||
15612 TATCAGTAGAGGTGAAGAGGTAAAGGACACATGCCAAGATTAAATCAAGATTAATA 15553
Qy 899 GCGCACTTAGCCAGTGGGGAAACAGACTCACACACAGTCCACAGCCAGAGTGGTCAGG 958
Db |||||
15552 TTGAAGGTAATTGAGGTTTTTGTCTCGGGGCACGCCACTAGTTAAGTAATATAGTTGACA 15493
Qy 959 GCCAAGATGGGAGACACGGGGAG-----AAAGGTGAGGTGGGATGGGAGGGG 1008
Db |||||
15492 GGCAGTTTGGAGATAAGATCTGTTGAATTTCTTCTGTTAGGCTTTCTGAGGATTC 15433

```

```
QY 1009 TCAGGCAAGAGGGTCAGGGCCAGCGTCGAGGGAAGCCCTGGGACTGTAGGAATTTAGAG 1068
Db 15432 TCAGTGGAAATTTTGTAAAGGTTATAGTTTGTGGTGGAAATCTGGTTTACGAATCGTG-G 15374
QY 1069 GAGGTACCTGACCGGCATGTTTGGTGGAGGAGATTCAAGAACTCTTCCCTGGAAGAGAGG 1128
Db 15373 TAGCTGAAGCCAGAAACAAAGAAATAATCTTGATCAAGGGTATGTAACTCATGAGGAAGG 15314
QY 1129 CTGTCGGAGCTGAGACTCATAAGATGAGTGGGAGGGTGTCC--AGGCAGAAAAGACCA 1185
Db 15313 AAAAGAAGCTTAAATGAGAAAGGTGGATAAACAGGAGAGCAAGATGGAAGAGCATT 15254
QY 1186 GCACCTACAAAAGCATGACTTTGAGAGAGCAATTCATCCATTCAACTGATGATTTTCAG 1245
Db 15253 TTTTGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 15194
QY 1246 ACTGGGCACGCTGGCTCATGCTGTAAATCCAGCACTTTTGAAGGCTGAATGGGAGGAT 1305
Db 15193 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 15134
QY 1306 GACTTGAGCCTAGGCATTTGTGACAAAGCTGGGCAACATGGTGAGAGCCCTGCCCTCACAA 1365
Db 15133 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 15074
QY 1366 AACAACAAACAAACAAACAAATCATATACCTGGTACCATGGTACAGGTACATAGAAA 1425
Db 15073 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 15014
QY 1426 TGACTCAGGCAGATATGGTGTCTCTCTACTGTGGGAGGGGCTTATATCGCAGTA 1485
Db 15013 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 14954
QY 1486 AGCAATAGAGGAGGGAATATAATCCTAAATGAGAGGTACAGATTTGAGACAAACAC 1545
Db 14953 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 14894
QY 1546 AGGCACAGGCATATGTACAGGGTAAAGAGGAAATCAGGGAAGGCTTCTCAGAGAAGT 1605
Db 14893 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 14834
QY 1606 GACATTTAAGCCGGGACATGAAGGATGAACGATTTAGTTCAACAAGGATGGGATGAAAG 1665
Db 14833 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 14774
QY 1666 GGGTGAGGTGATGGAGGCGAGAGGGAACCTCGAGATCATAGGCCTAGACAGGGGATCCTG 1725
Db 14773 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 14714
QY 1726 ACGCCCTTGAGGAAGTGAGAGAGAGACAGCGCAGTCGTAGTGGGTTAAGTAACAAAGCTG 1785
Db 14713 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 14654
QY 1786 AGAAGCAGGAAATCCCTGTCATGCAAGGCTGTGAGTCAGCTCAGAGTGTGTGGGCT 1845
Db 14653 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 14594
QY 1846 TTTGTTTTCTGGGAGCAGTCGATTTTAAGCAGGGAACAGCTGTATTTCAGAGTTGGGAA 1905
Db 14593 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 14534
QY 1906 GATCCTCTGCTGCTGCTGAAGGGATGAACCTGGAGGCTAGGAGCCAGGCTGATAGG 1965
Db 14533 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 14474
QY 1966 GAGGATCCAGGCTGATGGGAGGCTGGAGGTCGCGGTGATGACAGGGCTGGGGCCA 2025
Db 14473 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 14414
QY 2026 GGGATCGGAGGAAGGATGTAATTTGGGAGAGGCTTGGGCTCTGGCCGAGGAATGGATGG 2085
Db 14413 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 14354
```

## RESULT 5

US-10-242-515-3354/c

; Sequence 3354, Application US/10242515

; Publication No. US20040009488A1

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

; FILE REFERENCE: PC005C1

; CURRENT APPLICATION NUMBER: US/10/242,515

; CURRENT FILING DATE: 2002-09-13

; PRIOR APPLICATION NUMBER: 09/764,877

; PRIOR FILING DATE: 2001-01-17

; PRIOR APPLICATION NUMBER: 60/179,065

; PRIOR FILING DATE: 2000-01-31

; PRIOR APPLICATION NUMBER: 60/180,628

; PRIOR FILING DATE: 2000-02-04

; PRIOR APPLICATION NUMBER: 60/214,886

; PRIOR FILING DATE: 2000-06-28

; PRIOR APPLICATION NUMBER: 60/217,487

; PRIOR FILING DATE: 2000-07-11

; PRIOR APPLICATION NUMBER: 60/225,758

; PRIOR FILING DATE: 2000-08-14

; PRIOR APPLICATION NUMBER: 60/220,963

; PRIOR FILING DATE: 2000-07-26

; PRIOR APPLICATION NUMBER: 60/217,496

; PRIOR FILING DATE: 2000-07-11

; PRIOR APPLICATION NUMBER: 60/225,447

; PRIOR FILING DATE: 2000-08-14



```
; PRIOR APPLICATION NUMBER: 60/218,290
; PRIOR FILING DATE: 2000-07-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3354
; LENGTH: 16747
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (14037)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (14038)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (14039)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (14040)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (14041)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (14042)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (14043)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (14044)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (14045)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (14046)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (14047)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (14048)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (14049)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (14050)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (14051)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (14052)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (14053)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (14054)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (14055)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (14056)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (14057)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (14058)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (14059)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (14060)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (14061)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (14062)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (14063)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (14064)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (14065)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (14066)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (14067)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (14068)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (14069)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (14070)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
```

```
; NAME/KEY: misc_feature
; LOCATION: (14071)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (14072)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (14073)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (14074)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (14075)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (14076)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (14077)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:

Query Match      8.4%; Score 253; DB 16; Length 16747;
Best Local Similarity 27.2%; Pred. No. 1.9e-65;
Matches 659; Conservative 0; Mismatches 1740; Indels 22; Gaps 6;

QY 301 ACAGTGGCCCATCAGCGGGTGCAGTGGCTAGTGCCTGAAATCCCAACACTTTTGGGAGCG 360
DB 16152 AAAATGAGTTAGGCTGGCGGGTGGCTCAGGCTTAAATCCCAACACTTTTGGGAGCGC 16093

QY 361 GAGGTGGGTGGATCACTTGAAGTTCAGAGTTCAGACCGCCCTGGCCAAACATGGTGAAC 420
DB 16092 GAGGCTGGTGGATCACTTGAAGTTCAGAGTTCAGACCGCCCTGGCCAAACATGGTGAAC 16033

QY 421 CTGTCTCTACTAAAAATA-TAAAAATTAGCTGGCATGGTGGTACCTGTAATCCC 479
DB 16032 TCTGTCTCTACTAAAAATTACAAAAATTAGCTGGCATGGTGGTACCTGTAATCCC 15973

QY 480 AGATACTTGGGAGGTTGAGCGAGGAATCGCTTGAACCCGGAGGAGCAGATGTTGCAGTG 539
DB 15972 AGTACTTGGGAGGCTGAGCGAGGAATTCGCTGAACCCAGAGATGGAGGTTGTAGTG 15913

QY 540 AACCAAGACTGTGCCACTGCCTCAGTCCAGTTCGGGACAC-AAGTGAAGCCCTCCATCTCAAC 598
DB 15912 AGCCAGACTGGCCACTGCCTCAGTCCAGTTCGGGACACAGAGCGGAACCTCCGCTCAAAA 15853

QY 599 AAACAAACAAAAGCAGTGCCTCATGTAGGATTGAGTGGTGGAGGAGTGGAGCTG 658
DB 15852 AAACAAACAAAATTGGAGGAGATGAGATTGTAATGTAAATATATAGACATCATTTTA 15793

QY 659 TTGCAAGTGAAGCACTCACTAATACCAAGGTTGTAGTATCAGTGATTAACCATCAATGAT 718
DB 15792 TAGGAAATAGTGAAGTCTGAAAGACCTCTATGATGGGTTAAATGGGGAAGATGTTCAAT 15733

QY 719 CCAGGTAAAGCCCTGAGGTTTCAAGAAAGATGCGGAGCGCTTCAAGGTGCTGGGATG 778
DB 15732 TCTAGTCAAGAAAATATTTACAGACAAATTTTCTGTCAAGTTCAGAGGATGGAGAACAGA 15673

QY 779 GTGGGCAAGCCCTCGAATAATAGAAACAGTTCTCTGTATTACAAACAGAAAGCAGGAGGCC 838
DB 15672 ATGGAGCGAAGAGGCAATTGAGAAAGTGGTGGCATTAATTCAGGTGAGAGGTAAAGACAG 15613

QY 839 CATGCTGGGTGCTGCCAGGAACCTCAGTGAAGTAACTAAGACAGACCGCGTGTCTCCCA 898
DB 15612 TATCAGTAGAGGTGAAAGAGGTAAAGGACACATGCCAAGATTAAATCAAGGATAAATA 15553
```

```
899 GCGCACCTAGGCCAGTGGGGAACAGACTCACCACAGCTCCAGCCAGAGTGTGTCAGG 958
1552 TTGAAGGTAATTGAGGTTTTTCTCGGGGCAAGCCACTAGTTAAGTATATATAGTTGGACA 15493
959 GCCAAGATGGGGAACAGCAGGGGAG-----AAAGTCAAGGTGGGATGGGAGGGG 1008
15492 GGCAGTTTGGAGATAAGATCCTGTTGAATTTCTTGTGTTTAAAGGCTTCTGAGGATTC 15433
1009 TCAGGCAAGAGGGGTCAAGGCCAGGCTGAGGGAAGCCCTGGGACTGTAGGAATTTAGAG 1068
15432 TCAGTGGAAAATTTTCTAAGGTTATAGTTTGGGTGGAATCTGGTTTACGAATCGTG-G 15374
1069 GAGGTACCTGACCCGGCATGTTTGGTGAGGGAGATTCAAGAAAGTCTTCTCGAAGAGAGG 1128
15373 TAGCTGAAGCCAGAACAAAGATAATCTTGATCAAGGGTATGTAACTCATGAGGAAGG 15314
1129 CTGTCGGAGCTGAGACTCATAAAGATGAGTGGGAGGAGGTTTCC---AGCGAAGAACCA 1185
15313 AAAAAAGAGCTTAAATGAGAAAGGCTGGATAAACAAGAGAGACAAGATGGAAGAGCATT 15254
1186 GCACCTACAAAAGCATGACTTTTGAGAGAGGATTCATCCATTCACATGATGAATTTTCAG 1245
15253 TTTTTTTTTTTTTTTGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 15194
1246 ACTGGGCGAGCTGGCTCATGCTGTAATCCAGCACTTTGGAAGGCTGAATGGGAGGAT 1305
15193 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 15134
1306 GACTTGAGCTAGGCATTTGTGACAAAGCCTGGGCAACATGGTGAGACCTGCCTCCACAA 1365
15133 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 15074
1366 AACAACAACAACAACAATAATATACCTGGTACCATGGTACCAGGTACATAGAAA 1425
15073 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 15014
1426 TCACTCAGCAGATATGGTGTCTCTCTACTGTGGAGAGCGGCTTTATATCTGAGTA 1485
15013 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 14954
1486 AGACAATAGAGGAGGGAATATAATCTTAAATAGAGAGTAAGATTTGAGAGCAACAC 1545
14953 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 14894
1546 AGGCGACAGGCATATGTCAGGGTAAAGAGGAATCAGGGAAGGCTTCTCAGAGAGGT 1605
14893 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 14834
1606 GACATTTAAGCCGGGACATGAAGGATGAACGAGTTAGTTCA CCAAGGATGGGATGGAAG 1665
14833 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 14774
1666 GGGTGAGAGTGTAGGAGGAGAGGAACTGCGAGGATCATAGGCTTAGAGAGGAGTCTCTG 1725
14773 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 14714
1726 AGCCCTTTCAGGAGTGTAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1785
14713 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 14654
1786 AGAAGCAGGAGAAATCCCTGGTCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1845
14653 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 14594
1846 TTTGTTTTCTGGGAGAGTGCATTTTAAAGCAGGGAACAGCTGTATTTCAGAGTTGGGAA 1905
14593 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 14534
1906 GATCTGTGTTCTGCTCCCTGAAGGAGTGAACCTGAGGAGTGAAGGAGGAGGAGGAGG 1965
14533 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 14474
1966 GAGGATCCAGGAGTGTAGGAGGAGGCTGGGAGGCTCCGCGGTGATGGAACAGGAGTGGG 2025
```





Db 56815 ACAAATTTTTTACAAAAATAAATAAGTAATAATAAAGTCCCTATGCGCTAGTTGA 56756  
QY 2011 CCAGGGCTGGGCCAGGGATGGGAGAGAAAGAGTAATTGGAGAGGGCTGGGGCTCTGG 2070  
Db 56755 GAAGTGAAGATTATAGAGCTCTAATGATGTTAGATTTAGAGACATTAAGAAGGTTAGGGC 56696  
QY 2071 CCGAGGAATGGATGGTGGCTGAAACAGGAGAGAGAGATGCTTAGCCACTTTGGNAC 2130  
Db 56695 CGGCGATGGTCTACGCCCTGTATCCAGCACTTTGGGAGGCTTGAGCGAGGTAGA --- 56639  
QY 2131 ACAGTAGGCAAGGACAGGACACCCCAAGGGGAAGTCCCAAGAGACACGACAGGCTG 2190  
Db 56638 TCAGAGGTGAGGATGGAGACAGCCTGCGCAAGATGTTGAACCCCGTCTCTACTAA 56579  
QY 2191 GCATTGACAGGGAAGTCTCTCTGAGCAGGTGCTTGGATAGAGGAGGAATAATGTGC 2250  
Db 56578 AAATACAAAAATTAGCTGGGTGCTGGCAGGTGCTGTAATCCAGCTACTTTGGGAGGC 56519  
QY 2251 AGTTCCATCTCTCCCTCTCTCTGTTCAACCTTAACTACATGGGACAGACCCAGTG 2310  
Db 56518 CGAGGCAAGAGATCGC-----TTGAAACCGGAGTTGGAGGTTCAGTACCTGAGAT 56465  
QY 2311 GGACTCCATAAATGATGGGTGGTGGATGGAAGGAAGGAGGAGAAACAACCTTTCAT 2370  
Db 56464 CGCTCCACTGCACCTAGCTTGGG-----CAACAAGCAAGACTCTGT 56422  
QY 2371 TCATCTGGTTATTTACAGAACAGGCGAGGTGCGGTGCTACGTTGCCATTTAGCACT 2430  
Db 56421 CTCAAAAACAATAAATAAAGGCTGGGCACAGTGGTTATGCTCTGTAATCCAGCACT 56362  
QY 2431 TTGGGAGGCTGAGTGGGTGATTTACCTCAGGTGAGGATTCAGACAGCCTAGACAC 2490  
Db 56361 TTGAGAGCCGAGAGGCGGATCACCCTGAGGTGCGGAGTTTGAGACCGCCTGACCAAC 56302  
QY 2491 GTAGAGAAACCCATCTCTACTGAAGATATAAATAAGTGGGCTAGTGGCATATGCT 2550  
Db 56301 ATGGAGAACTCGCTCTCTACTAATAAATAAATAATTCGCGAGTGTGGTATGCTCT 56242  
QY 2551 GTATCCAGCTAGTGGGAGCTGAGGAGGAGGAGTAATCGTTGAACCCGAGAGCAGG 2610  
Db 56241 GTATCCAGCTATTCGGGAGGCTGAGGCAAGAGATCGCTTGAACCCGAGGCGGAGG 56182  
QY 2611 TTGGGTGAGCTGAGATCGTGGCAATTCGCTCCAGCTGGGTGACAAAGCAAGCTCGT 2670  
Db 56181 TTGAGTGAACCGAGATCGCGCAATTCGCTCCAGCTGGGCAATAAGATGAAGAACTCCA 56122  
QY 2671 CTCATAATAATAATAATTACAAACAGAA 2700  
Db 56121 TCTGAAAAAATAAATAAATAAATAA 56092

RESULT 9  
US-10-412-277-3/c  
; Sequence 3, Application US/10412277  
; Publication No. US20030175791A1  
; GENERAL INFORMATION:  
; APPLICANT: GUEGLER, Karl et al  
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC  
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES  
; FILE REFERENCE: CL001067DIV  
; CURRENT APPLICATION NUMBER: US/10/412,277  
; CURRENT FILING DATE: 2003-04-14  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 786431  
; TYPE: DNA  
; ORGANISM: Human  
; NAME/KEY: misc feature  
; LOCATION: (1)...(786431)

OTHER INFORMATION: n = A, T, C or G  
US-10-412-277-3  
Query Match 8.2%; Score 245.8; DB 15; Length 786431;  
Best Local Similarity 84.2%; Pred. No. 2.3e-62;  
Matches 277; Conservative 0; Mismatches 52; Indels 0; Gaps 0;  
QY 282 TCTGTCAAATGGAGTGATTAACAGTGCCTCATCAGCGGGTGCAGTGGCTAGTCCCTGAAT 341  
Db 572672 TTTTTCATTACTCTATTAACAGTGACTAGAGGCGAGGTGCTGGCTCATGCTATAT 572613  
QY 342 CCCAACACTTTGGGAGGCGGAGGTGGTGGATCACTTTGAGTCAAGGATTCGAGACCGC 401  
Db 572612 CCCAGCACTTTGGGAGGCGGAGGTGGTGGATCACTTTGAGTTCGGAGTTCCGAGACCGC 572553  
QY 402 CTGCGCAACATGTGAAACCCCTGCTCTACTAAAAATATAAAATTAAGCTGGCATGTG 461  
Db 572552 CTGCGCAACATGTGAAACCCCTGCTCTACTAAAAATATAAAATTAAGCTGGCATGTG 572493  
QY 462 GTGGTACTCTGTAATCCAGATACTTGGAGGTTGAGGAGGAGAAATGCTTTGAACCCGG 521  
Db 572492 GCGTACACCTGTGATCCAGCTACTTGGGAGGTTGAGGAGGATAATCACTTTGAACCTGG 572433  
QY 522 GAGGACATGTTGAGTGAACCAAGACTGTGCCACTGCCTCCAGTCTGGSCAACAGAT 581  
Db 572432 GAGTGGAGGTTGAGTGAACCAAGACTGTGCCACTGCCTCCAGTCTGGGTCAGAGT 572373  
QY 582 GAGCTCCATCTCAAAACAAACAAACAAA 610  
Db 572372 GAACTCCGCTCAAAACAAACAAACAAA 572344

RESULT 10  
US-09-997-722-154/c  
; Sequence 154, Application US/09997722  
; Publication No. US20040072154A1  
; GENERAL INFORMATION:  
; APPLICANT: Morris, David  
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER  
; FILE REFERENCE: A-71171/RMS/DCF  
; CURRENT APPLICATION NUMBER: US/09/997,722  
; CURRENT FILING DATE: 2001-11-30  
; PRIOR APPLICATION NUMBER: US 09/747,377  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: US 09/798,586  
; PRIOR FILING DATE: 2001-03-02  
; NUMBER OF SEQ ID NOS: 301  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 154  
; LENGTH: 96594  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-997-722-154

Query Match 8.2%; Score 245.6; DB 12; Length 96594;  
Best Local Similarity 48.2%; Pred. No. 8.7e-63;  
Matches 1158; Conservative 0; Mismatches 1174; Indels 68; Gaps 14;  
QY 314 GCGCGTGCAGTGGCTAGTGCCTGAATCCCACTTTGGGAGGCGGAGGTGGTGGAT 373  
Db 17837 GCGCGGTGGTGGCTTATGCTGTAAATCCAGCACTTTGGGAGGCGGAGCGGAGAT 17778  
QY 374 CACTTGAGGTGAGGATTCGAGACCACTGGCCAACTGGTGAACCCCTGCTCTACTA 433  
Db 17777 CACTTGAGGTGAGGATTCGAGACCACTGGCCAACTGGTGAACCCCTGCTCTACTA 17718  
QY 434 AAAATATAAAATTAAGTGGCATGGTGGTGGTGAATCCAGATCTCTTAATCCAGATCTTGGGAGG 493  
Db 17717 AAAATATAAAATTAAGCAGGATGATGTGACACCTATTAACCCAGCTATTTCAGTAGG 17658  
QY 494 TTGAGCAGGAGAAATCGCTTGAACCCGAGGAGGAGATGTGAGTGAACCAAGCTGTGC 553

Db 17657 CTGAGACATGAGAATCTCTTGAACCCAGCTGGTAGAGGTTGCAGTGAGCCCAAGATCATAC 17598  
Qy 554 CACTGCACCTCCAGTCTGGGCAACAGAGTGAGCCTCCATCTCAAAACAAACAAAAAGC 613  
Db 17597 CACTGCACCTCCAGCCTGGGCGACAGAGTGAGACTCTCTCTCAAAAAACAAACAAAA--- 17541  
Qy 614 AGTGCCCATCATGTAGGATTGAGTGATTTGAGTGAGGACTGAGCCTTGTGCAAGTGAGCA 673  
Db 17540 -----CTTCTCTGGGTAGGCAAAAGGGGCTGGGGAGTGACTTAAATCACCAATGGCCAA 17487  
Qy 674 CTCACATAAC-ACCAGGTTGTAGTATCAGTGATAACCATCAATGATCAGGTAAGAGCCCT 732  
Db 17486 TTATTTATCATCATGTCTAGTTAGGAGGCTTACGTAAACCCAGNAGGATAGGGCT 17427  
Qy 733 GAGGTTTCAGAAAGATGCGGAGCCCTTTCAAGGTGCTGGGAATGGTGGGCAAGCCCTC 792  
Db 17426 CAGGAGCTTCTAGGTTGGTGAGAACATGTAGAGATGGTGGGAGGCTGGTGTGCCCTGGAG 17367  
Qy 793 GNAATAATAGAACAGTTCTCTGTATTACACAGAAAGCAGGAGGCCCATGCTGGGTGCTG 852  
Db 17366 AGGCGATGGAAGTGCCACACCCCTTCCACATACCTTGTCTGTGTATTCTTCCAGCTG 17307  
Qy 853 CCAGGAATCTCAGTAGTAACATAAGACAGCACCGGTGCTCTCCAGCGCACCTAGGCCA 912  
Db 17306 GCTGCTTCTGAGTTGTTTATATATAGCTGATTAATCTAGTAAGGAACCTGTTCTGTGACTG 17247  
Qy 913 GTGGGGAACAGACTCACACACAGTCCAGCCAGAGTGFTCAGGGCCAGATGGGAA 972  
Db 17246 AATGTGGAGGACTTACCCTTGTGCTTGTAGTCTGAAGTGGGGGCTTGGAAATGGGGTG 17187  
Qy 973 GCACGGGAGAAAGTTCAGGTTGGATGGGAGGGGTCAGGCGACAGAGGGGTCAGGGCCA 1032  
Db 17186 GACGGGAGGACTTAACATTTGGGCTCTGCACTAACCTGGCAGTTAGTGTACAGCTG 17127  
Qy 1033 GCCTGAGGGAAGCCCTGGGACTGTAGGAATTTAGAGGAGGTACCTGACCCGGCATGTTTG 1092  
Db 17126 AGCTAAATGTAGGACACCCAGTGTCTGCTAAAGACTGAATTTGTGGGAAA----- 17073  
Qy 1093 GTGAGGAGATTCAGGAAGTCTTCTGGAGAGAGGCTCTCGAGCTGAGACTCATAGA 1152  
Db 17072 ---AAAAACACCACACATTTGGTGACAGAAAGTGAGGAATTTATATATGGGTACAGAGTAG 17016  
Qy 1153 TGAGTGGGAGGAGTGTTCAGGCAGAAAGACCAGCACCTTACAAAAGCATGACTTTCAGAG 1212  
Db 17015 AGAGGAATGGAGTTTCTTTACAGTGGTATAGGCA-----TAGAATG 16973  
Qy 1213 AAGCATTTCAATCCATTCAGATGAATTTTCAGACTGGGACGCTGGCTCATGCCCTGTAA 1272  
Db 16972 AAATATTCAGCCTTTAAAGAAAAAAATTT--GACTAGATGAAGTGGCTCATGCTGTGTA 16915  
Qy 1273 TCCAGCACTTTGGAAAGGCTGAATGGGAGGATGACTTGAGCCTTAGGCATTTGTGACAAG 1332  
Db 16914 TCCAGCACTTTGGGAGGCCAAGGCGAGGATATCGCTTGAGCCCGAGGTTTGAGACCAG 16855  
Qy 1333 CCTGGSCAATGGTTCAGAGCC--TGCCCTCCAAAAACAAACAAACAAACAAATCAT 1390  
Db 16854 CCTAGGCAACAAAGCAAGCCCGATCTACAAAAATATAAAAGTTAGCCAGCATGGT 16795  
Qy 1391 TATACCTGGTACCATGGGTACAGGTACATAGAAATGACTCAGGCGAGATATGGTGTCTC 1450  
Db 16794 GGTACACTCTCTGTAGTCTCACCTTCTCGGAGGCTGAGGCAAGAGTATCATCTGAGCCCA 16735  
Qy 1451 TCTACTGTGGGAGAGCGGGCTTACTGTGAGTAGAAGATAGAGAGGAGGGAATATAAT 1510  
Db 16734 GGAGTTTAGAGGCTGCGATGAGCAACCATCATCTTACTGTCACTCCAGCCTGGATAGCTGAG 16675  
Qy 1511 CCTAAAATGAGAGGTACAGATTTGAGAGCAAAACACAGGGGCACAGGCAATGTACGAGGCT 1570  
Db 16674 CAAGACCTCTCTGMAAAATAATAATAATAACAATACAAATTAAGTTAAATAAATAAATA 16615  
Qy 1571 AAAAGGGGAATCAGGGAGGCTTCTCAGAGAGGCTGACATTTAAGC-----GGGACATG 1625  
Db 16614 AAAAAAAAATCCTGCAACATGAGATGACACAGATGAACCTTGAAGCCATTTAAGCCAAAGTG 16555

Qy 1626 AAGGATGAACGAGTTAGTTTACCAAGGATGGATGGAAGGGGTGAGAGTGATGAGGCA 1685  
Db 16554 AAGTAAGCCAGTACAGAGACAAATATCTACATGATTCCTACTGCAAGAGGTATCTAAATA 16495  
Qy 1686 GAGGAACTGAGGATCATAGGCCCTAGACAGGGGATCTCGACGCCCTTTGAGGAAGTGAGA 1745  
Db 16494 GGCAAACTCATAGAAGCAAGAGGGGCACAGTGGTTGCCAGGGTAGGGCTTTTCAGGAAA 16435  
Qy 1746 GAAGACAGCGAGCTGCTAGTGGGTTAAGTAAACAAGCTGAGAAAGCCAGGGAATCCCTG 1805  
Db 16434 TGGGAGTTGTTAATCAAAAAAGTATAAAGTTTCAGTAATGCAAAATGAACGCAATCTAGA 16375  
Qy 1806 GTCATGCAAGGCCCTGTGAGTCAAGTCAAGAGTGTGGGCTTTGGCTTTTCTCTGGGAGCAG 1865  
Db 16374 GATCTGCCATACAAATTTGTGC-TGTGTGTTACTATACATGATAAACTCTCTTAAAGG 16316  
Qy 1866 TCGATTTTAAAGCAGGAAACAGCTGTATTACAGATTTGGGAAGATCTGTGTGTTGCTGCTG 1925  
Db 16315 TTAAAGAAAAACAATTTAAAAAAAATTTTATAGAGACAAGGTCTTGTCTCTGCACTCAG 16256  
Qy 1926 AAGGGGATGAACCTGGAGCTAGGAGCCAGGGTGTATAGGAGGATCCAGGGTGATGGGG 1985  
Db 16255 GCTAGAGCGCAGTGGCACAATCATAGTTCACTGCAAGCCTCAGACTCTTAAGCTCAAGTGA 16196  
Qy 1986 AGGCTGGGAGGTCGCCGCTGATGGACCAGG---GCTGGGGCCAGGGGATGGGAGGAAGG 2042  
Db 16195 TCCTCTGCTCAGCCTCTGGNAGATCTGGAACTACAGGTGATGACATCATGCTCAACT 16136  
Qy 2043 AGTAATTTGGGAGAGCCTGGGCTCTGGCCGAGGAATG-----GATGTGGGCTGAAAC 2096  
Db 16135 AATTTTAAATTTTATAGAGACAAGTCTCCATTTGTGCAGACTGGCTCTTGAATCTC 16076  
Qy 2097 AGGAGAGGAGAGATGCTTAGCCCACTTTGGAAACACAGTAGGCAAGGACAGAGACACC 2156  
Db 16075 CTGGCTCAAGAGATCTTGTCTCCTCAGCCTCCCAAAGTGTGGGAATACAGAAAGTGAGC 16016  
Qy 2157 CAAGGGGAAGTCCCAAGAGACCAAGCAGGCTGGCATTTGGACAGGGAAGTGTGCTGTG 2216  
Db 16015 TACCATGCCAGCCAAATTAATCTATTAAGAAAAAAGAAAAAGAAACATAATCT 15956  
Qy 2217 AGCAGGTGTCTTGGATAAGGAGGAAAAATGGTGC-----AGTTCCATCTCTCTCCC 2267  
Db 2268 TCTCTGTTCAACCTCTAACTACATGGGCAAGGACCCAGTGGGACTCCATAAATGATG 2327  
Qy 15895 TATAGTATTTATCCCATTTACAGAGCTCAAAACAGGCAAACTGAGCTATATTATTAT- 15837  
Db 2328 GGATGGGTGGATGGAAGGAAGGAGGAAAAACAATCTTTCAATTCATCTCTGGTTATTTAC 2387  
Qy 15836 AGATGTATAGTACAGGCAAACTATGAACACAAGAAAGCAATTTATCAAAAAATTTGGAC 15777  
Db 2388 AGAACAGGCCAGGTGGGTGCTCAGCTTGGCAATTTAGACTTTGGGAGGCTGAGGTG 2447  
Qy 15776 AATAGGCGCAGGAGCGGTGGCTCATGCTGTAAATCCAGCACTTTGGGAGGCGAGGCGAG 15717  
Db 2448 GTGGATTACTCAGGTCAGGAGTTCAAGACCAAGCTAGACAACGTAGAGAAACCCCATCT 2507  
Qy 15716 CGAGATTACTGAGGTACAGAGTTTAAAGCCAGGCTGGCCCAACATGATGTAACCCCGTCT 15657  
Db 2508 CTACTCAAGAT--ATAAAATTAGCTGGGCTAGTGGCATATGCTCTAATTCACAGTAGT 2565  
Qy 15656 CTACTAAANTAGAAAAATTAGCCAGCCATGTTGGCAGGCGCTCTAATTCGAAGTACT 15597  
Db 2566 CGGGAAGCTGAGCAGGAGAAATCGCTTGAACCCGAGAGCAGAGGTTGCGGTGAGCTGAG 2625  
Qy 15596 TGGGAGCTGAGCAGGAGAAATGCTTGAACCCAGGAAGCAGAGGTTGAGTGAGCGCGAG 15537  
Db 2626 ATGCTGCCATTCACCTCCAGCCTGGGTGACAA-AGCAAGACCTCTGCTCTCAATAAATAA 2684  
Qy 15536 ATGCGGCCACTGCACTCCAGAAATGGGCAACAGAGCAAAACTCCGCTCTCAAAATAAATA 15477



```
RESULT 11
US-10-312-495-13/c
; Sequence 13, Application US/10312495
; Publication No. US20030165495A1
; GENERAL INFORMATION:
; APPLICANT: BIOGEN, INC.
; TITLE OF INVENTION: NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: A099PCT000454-124
; CURRENT APPLICATION NUMBER: US/10/312,495
; CURRENT FILING DATE: 2002-12-23
; PRIOR APPLICATION NUMBER: 60/213,630
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 60/283,813
; PRIOR FILING DATE: 2001-04-13
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 13
; LENGTH: 10500
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-312-495-13

Query Match      8.2%; Score 244.8; DB 15; Length 10500;
Best Local Similarity 85.3%; Pred. No. 4.8e-63;
Matches 273; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 302 CAGTGCCTCATCAGCCGGGTGACGTGGCTAGTGGCTGAAATCCCAACACTTTGGAGGCGG 361
Db 2586 CTGGGTCTACTAACTACAAAATTTAGCTGGGCATGGTGGCGGTGCTGTATCCAGCACATTTGGGAGCGG 2527
QY 362 AGTGGGTGATCAGTTCAGAGTTCAGAGCCAGCGCTGGCCAAACATGTGAAACC 421
Db 2526 AGCGGGGTGATCAGTTCAGAGTTCAGAGTTCAGAGCCAGCGCTGGCCAAACATGTGAAACC 2467
QY 422 CTGTCTCTACTAAAAATATAAAATTTAGCTGGGCATGGTGGCTGCTGACCTGTATCCCGAG 481
Db 2466 CTGTCTCTACTAACTACAAAATTTAGCTGGGCATGGTGGCGGTGCTGTATCCCGAG 2407
QY 482 ATACTTTGGGAGTTGAGGCGAGGAGATCGTTGAAACCGGAGGAGAGATTTGCAGTGAA 541
Db 2406 CTACTTTGGGAGGTTGAGGCGAGGAGATCGTTGAAACCGGAGGAGGTTGCAGTGAC 2347
QY 542 CAAAGACTGTGCCTGCACTCCAGTCTGGGCAACAGAGTGAGCTCCATCTCAACAAA 601
Db 2346 CTCAGATCGTGCCTGCACTCCAGTCTGGGCAACAGAGTGAGCTCCATCTCAACAAA 2287
QY 602 CAAACAAAAGCAGTGCCCA 621
Db 2286 ACAAAAACAGGCGGGTGCA 2267

RESULT 12
US-10-027-632-256849/c
; Sequence 256849, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 256849
; LENGTH: 546
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-256849
```

```
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 256849
; LENGTH: 546
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-256849

Query Match      8.1%; Score 243; DB 13; Length 546;
Best Local Similarity 84.0%; Pred. No. 3.6e-63;
Matches 273; Conservative 1; Mismatches 51; Indels 0; Gaps 0;

QY 287 CAAATGGAGTGATAACAGATGCCCATCAGCCGGGTGAGTGGCTAGTGCCTGAAATCCCAA 346
Db 541 CACAAGGGGTTTAAACACAGCTATAGGCGCGGAGCCATGGCTCAGCCTGTATCCAG 482
QY 347 CACTTTGGGAGGCGGAGGTGGGTGATCATCTGAGGTTCAGAGTTCGAGACCAAGCTGGC 406
Db 481 CACTTTGGGAGGCTGAGTGGGTGATCATCTGAGGTTCAGAGTTCGAGACCAAGCTAGC 422
QY 407 CAACATGGTGAACCCCTGCTCTACTAAAAATATAAAATTTAGCTGGGCATGGTGGCG 466
Db 421 CAACATGGTGAACCCCTGCTCTACTAAAAATATAAAATTTAGCTGGGCATGGTGGCGCA 362
QY 467 TACCTGTAATCCAGATACCTTGGGAGGTTGAGGCGAGAGATCGCTTGAACCCGGGAGGC 526
Db 361 TGCTGTAATCCAGATATTTGGGAGGCTGAGGCGAGAGATCGCTTGAACCCAGAGGC 302
QY 527 AGATGTGAGTGAACCAAGACTGTGCCACTGTCCACTCCAGTCTGGGCAACAGAGTGAGCC 586
Db 301 AGAGGTTGAGTGAACCAAGATTTACGCCACTGCCACTCCAGCTGGGCAACAGAGCAAGAC 242
QY 587 TCCATCTCAACAAACAAACAAA 611
Db 241 TCCATCTCAACAAACAAACAAA 217

RESULT 13
US-10-027-632-256849/c
; Sequence 256849, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 256849
; LENGTH: 546
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-256849
```



Db	884	GCAGAGAGTCGAGATCGTGCCACTCCAGCTGGGCAACAGAGTCGAGACTCTGTCT	943
Qy	594	CAAAACAACAAACAAAA	611
Db	944	CAAAAAAAAAAAAAAAAA	961

Search completed: September 18, 2004, 01:40:19  
Job time : 1348.64 secs

***This Page Blank (uspto)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 17, 2004, 07:12:45 ; Search time 7312.95 Seconds  
(without alignments)  
12254.481 Million cell updates/sec

Title: US-10-669-693-3\_COPY\_1980\_4980

Perfect score: 3001

Sequence: 1 ctttctctctccctaccccc.....atcctggttaacacagtga 3001

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: em\_estba:\*

2: em\_esthum:\*

3: em\_estin:\*

4: em\_estnu:\*

5: em\_estov:\*

6: em\_estpl:\*

7: em\_estro:\*

8: em\_htc:\*

9: gb\_esti:\*

10: gb\_est2:\*

11: gb\_htc:\*

12: gb\_est3:\*

13: gb\_est4:\*

14: gb\_est5:\*

15: em\_estfun:\*

16: em\_estom:\*

17: em\_gss\_hum:\*

18: em\_gss\_inv:\*

19: em\_gss\_pln:\*

20: em\_gss\_vrt:\*

21: em\_gss\_fun:\*

22: em\_gss\_man:\*

23: em\_gss\_mus:\*

24: em\_gss\_pro:\*

25: em\_gss\_rod:\*

26: em\_gss\_phg:\*

27: em\_gss\_vrl:\*

28: gb\_gss1:\*

29: gb\_gss2:\*

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	241.8	8.1	757	28	AQ528478
C 2	241.6	8.1	760	28	AL691744
C 3	240.2	8.0	478	28	B67141
4	240	8.0	808	28	BZ611126

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

C	5	239	8.0	833	28	BZ599987
	6	237.6	7.9	553	14	CA438009
	7	237.4	7.9	551	12	BM754642
C	8	237.4	7.9	657	14	CF123536
C	9	237.2	7.9	696	13	BU616112
	10	237	7.9	690	28	AQ415537
	11	236.8	7.9	494	10	AW471332
C	12	236.8	7.9	932	13	BQ860197
	13	236	7.9	928	28	AQ45514
	14	236	7.9	656	29	AG035209
	15	235.8	7.9	466	28	AQ590440
C	16	235.6	7.9	470	28	AQ226326
	17	235.6	7.9	588	13	BA488003
	18	235.6	7.9	677	29	AG169901
	19	235.4	7.8	412	10	BE062478
	20	235.2	7.8	490	10	BE281645
	21	235.2	7.8	746	12	BG774140
C	22	235	7.8	560	10	BE155951
C	23	235	7.8	724	14	CB555645
C	24	235	7.8	995	12	BM423099
	25	234.8	7.8	477	28	AQ221138
	26	234.8	7.8	611	14	CD704731
C	27	234.6	7.8	501	13	BA485916
	28	234.4	7.8	672	13	BA501311
	29	234.2	7.8	675	13	BU664420
C	30	234.2	7.8	3010	11	BC032827
	31	234	7.8	416	10	BE062476
C	32	234	7.8	614	28	AQ627870
C	33	234	7.8	678	28	AZ518835
C	34	234	7.8	687	14	CA448731
	35	233.6	7.8	588	28	AQ45949
	36	233.6	7.8	877	28	AQ739838
C	37	233.2	7.8	521	13	BA485089
	38	233.2	7.8	642	28	B59854
	39	233.2	7.8	648	28	AQ266668
	40	233.2	7.8	648	28	BZ611349
	41	233.2	7.8	685	28	AQ423897
	42	232.8	7.8	439	9	AL712324
C	43	232.4	7.7	439	9	AA515728
C	44	232.4	7.7	463	28	AQ535344
	45	232.4	7.7	541	28	BZ604634

## ALIGNMENTS

RESULT 1	AQ528478	757 bp	DNA	linear	GSS 18-MAY-1999
LOCUS	RPCI-11-349N11-TV	RPCI-11	Homo sapiens	genomic clone	
DEFINITION	RPCI-11-349N11, genomic survey sequence.				
ACCESSION	AQ528478				
VERSION	AQ528478.1	GI:4840591			
KEYWORDS	GSS.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	1 (bases 1 to 757)				
AUTHORS	Zhao, S., Adams, M.D., Nierman, W., Malek, J., de Jong, P. and Venter, J.C.				
TITLE	Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready Map Building				
JOURNAL	Unpublished (1997)				
COMMENT	Other GSSs: RPCI-11-349N11.TJ Contact: Shaying Zhao, William Nierman, Mark Adams Department of Eukaryotic Genomics The Institute for Genomic Research 9712 Medical Center Dr., Rockville, MD 20850 Tel: 301 838 0200 Fax: 301 838 0208 Email: hbeetigr.org Clones are derived from the human BAC library RPCI-11. For BAC				

library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (<http://bacpac.med.buffalo.edu/ordering>) or from Research Genet cs ([info@resgen.com](mailto:info@resgen.com)). BAC end search page: [http://www.tigr.org/tdb/hungen/bac\\_end\\_search/bac\\_end\\_search.html](http://www.tigr.org/tdb/hungen/bac_end_search/bac_end_search.html).  
 Seq primer: T7  
 Class: BAC ends.

#### FEATURES

location/Qualifiers  
 1..757  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="GDB:7633954"  
 /db\_xref="taxon:9606"  
 /clone="RPCi-11-349N11"  
 /sex="Male"  
 /cell\_type="Lymphocytes"  
 /clone\_lib="RPCi-11"  
 /note="Vector: pBACe3.6; Site\_1: EcoRI; Site\_2: EcoRI; RPCi11 Human Male BAC Library"

#### ORIGIN

```

Query Match      8.1%; Score 241.8; DB 28; Length 757;
Best Local Similarity 87.7%; Pred. No. 2.7e-43;
Matches 264; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 314 GCCGGTCAGTGCGTGTGCTGAAATCCACACACTTTGGAGCGGAGGTGGGTGAT 373
DB 268 GCCTGGTCAGTGGCTGATGCTGTAATCTCAGCACATTTGGAGGCTGAGGTGGTGGAT 327
QY 374 CACTTGAGTTCAGGAGTTTCAGACAGCAGCTGGCCAAACATGGTGAACCTGTCTCTACTA 433
DB 328 CACTTGAGTTCAGGAGTTTCAGACAGCAGCTGGCCAAACATGGTGAACCTGTCTTACTA 387
QY 434 AAAATATATAAATAGCTGGGATGGTGGTGGTACCTGTATCCAGATATCTGGGAGG 493
DB 388 AAATATACAAAATAGCTGGGATGGTGGCAGGCACCTGTATCCACCTACTTGGGAGG 447
QY 494 TTGAGGAGGAGATCGCTTGAACCCGGAGGCAGATGTGCAGTGAACCAAGACTGTGC 553
DB 448 CCGAGGAGGAGATTTGCTGAACCCGGTGGAGGTTCAGTGAAGGTCGAC 507
QY 554 CACTGCATCCAGTCTGGGCAACAGAGTGAGCTCCATCTCAACAAACAAACAAAGC 613
DB 508 CACTGCATCCAGCTGGGTACAGATGAGACTCTGTCTCAAAAAAANAANAAGA 567
QY 614 A 614
DB 568 A 568

```

#### RESULT 2

AL691744/c  
 LOCUS B67141 760 bp mRNA linear EST 04-SEP-2003  
 DEFINITION DKFZp313M0830.s1.313 (synonym: hlcc2) Homo sapiens cDNA clone genomic survey sequence.  
 ACCESSION AL691744  
 VERSION AL691744.1 GI:19617321  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 760)  
 AUTHORS Koehrer, K., Beyer, A., Mewes, W., Weil, B. and Wiemann, S.  
 EST (Koehrer, K., Beyer, A., Mewes, W., Weil, B. and Wiemann, S.)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: MIPS

MIPS  
 Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany  
 This is the 3' sequence of the clone insert  
 Cloned from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email [s.wiemann@dkfz-heidelberg.de](mailto:s.wiemann@dkfz-heidelberg.de);  
 sequenced by BMFZ (Biomedical Research Center at the Charite,

Berlin/Germany) within the cDNA sequencing consortium of the German Genome Project.

ri sequence also available.  
 This clone (DKFZp313M0830) is available at the RZPD in Berlin.  
 Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: [clone@rzpd.de](mailto:clone@rzpd.de).

#### FEATURES

location/Qualifiers  
 1..760  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="DKFZp313M0830"  
 /dev\_stage="adult"  
 /lab\_host="DH10B"  
 /clone\_lib="313 (synonym: hlcc2)"  
 /note="Vector: piriplex2; Site\_1: SfiI; Site\_2: SfiI; cdna-collection"

#### ORIGIN

```

Query Match      8.1%; Score 241.6; DB 9; Length 760;
Best Local Similarity 85.9%; Pred. No. 3e-43; Indels 0; Gaps 0;
Matches 268; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 300 AACAGTGCCTCAGCCGGTGCAGTGTGCTAGTGCCTGAAATCCCAACACTTTGGAGGC 359
DB 323 AAAAGTCACCATAGGCTGGTGTGTGTTCCAGCCCGTAATCCAGCACATTTGGAGGC 264
QY 360 GGAGTGGTGGATCATCTTGAGGTTCAGAGTTCAGACAGCAGCTGGCCAAACATGTGAAA 419
DB 263 TGAGGAGGTGGATCACCTGAGGTCAAGAGTTCGAGACCCAGCTGGCCAAACATGTGAAA 204
QY 420 CCCTGTCTCTACTAAAAATATAAAATTTAGCTGGGCATGTGGTGCCTACCTGTAATCCC 479
DB 203 CCCCGTCTCTACTAGAAATACAAAACCTAGCTGGGTGTGGTGGCTGTGTAATCCC 144
QY 480 AGATACCTTGGGAGGTTCAGGAGGAGAAATCGCTTGAACCCGGAGGCAGATGTTCAGTG 539
DB 143 AGTACTTGGGAGGCTGAGACAGGAGAAATCTCTTGAACCCAGGAGAGATGTTCAGTG 84
QY 540 AACCAAGACTGTGCCACTGCACTCCAGTCTGGGCAACAGAGTGAGCCTCCATCTCAACA 599
DB 83 AGCCAAGATTTGGCACTGCATCCAGCTGAGCGAGAGTGAGACTCCATCTCAAAAA 24
QY 600 AACCAACAAAAA 611
DB 23 AAAAAAAAAA 12

```

#### RESULT 3

B67141/c  
 LOCUS B67141 478 bp DNA linear GSS 21-JUN-1998  
 DEFINITION CIT-HSP-2017022.TRB CIT-HSP Homo sapiens genomic clone 2017022, genomic survey sequence.  
 ACCESSION B67141  
 VERSION B67141.1 GI:2641119  
 KEYWORDS GSS.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 478)  
 AUTHORS Adams, M.D., Rounsley, S.D., Field, C.E., Bass, S., Linher, K., Golden, K., Berry, K., Granger, D., Suh, E., Wible, C., Shizuya, H., Simon, W. and Venter, J.C.  
 TITLE Use of a random BAC End Sequence Database for Sequence-Ready Map Building  
 JOURNAL Unpublished (1997)  
 COMMENT Other GSSs: CIT-HSP-2017022.TFB  
 Contact: Mark Adams  
 Department of Eukaryotic Genomics  
 The Institute for Genomic Research  
 9712 Medical Center Dr., Rockville, MD 20850, USA  
 Tel: 301 838 0200



Fax: 301 838 0208  
 Email: mdadams@tigr.org  
 Clones are available from Research Genetics (info@resgen.com). BAC  
 end search page:  
 http://www.tigr.org/tldb/hungen/bac\_end\_search/bac\_end\_search.html  
 Seq primer: M13 Reverse  
 Class: BAC ends.

## FEATURES

source  
 1. .478  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="GDB:7044072"  
 /db\_xref="taxon:9606"  
 /clone="2017022"  
 /sex="Male"  
 /cell\_type="Sperm"  
 /clone\_lib="ClT-HSP"  
 /notes="Vector: pBelOBAC11; Site\_1: HindIII; Site\_2:  
 HindIII"

## ORIGIN

Query Match 8.0%; Score 240.2; DB 28; Length 478;  
 Best Local Similarity 87.4%; Pred. No. 5.8e-43;  
 Matches 263; Conservative 0; Mismatches 38; Indels 0; Gaps 0;  
 QY 311 TCAGCCGGGTGCAGTGGCTAGTGCCTGAAATCCCAACACTTTGGGAGCGGAGGTGGTG 370  
 |||||  
 Db 398 TCAGCCAGGTACATTGGCTCATGCTGTATCCAGCACCTTTGGAGGCCAAGTGGGCG 339  
 QY 371 GATCACTTAGGTCAGAGTTCGAGACCCAGCTGGCCAAATGCTGAAACCCCTGTCTCTA 430  
 |||||  
 Db 338 GATCACTAGAGTTCAGGAGTTAGAGACCGAGCTGGCCAAATAGTGAACCCCTGTCTCTA 279  
 QY 431 CTAATAATATAAAATTTAGCTGGCGATGTTGCGTACCTGTAAATCCAGACTTTGGG 490  
 |||||  
 Db 278 CTAATAATATAAAATTTAGCCAGCATGTTGGGCTGTCCTGTAGTCCAGCTACTTTGGG 219  
 QY 491 AGGTTGAGGCGAGAGATTCGCTTTGAACCCGGGAGGAGATGTTGCAGTGAACCAAGACTG 550  
 |||||  
 Db 218 AGGCTGAGGCGAGAGATTCCTTTGACCTGGGAGGAGAGTTCGAGTGAACCAATG 159  
 QY 551 TGCCTAGTCACTCCAGTCTGGGCAACAGAGTGAAGCTTCCATCTCAACAAACAAACAAA 610  
 |||||  
 Db 158 CACCACTGCACTCCAGCTGGGCAACAGAGTGAAGCTCTCTCTCCAAACAAACAAAAG 99  
 QY 611 A 611  
 Db 98 A 98

RESULT 4  
 BZ611126 808 bp DNA linear GSS 08-JUN-2003  
 LOCUS WHAD87TF Human MCF7 breast cancer cell line library (MCF7\_1) Homo  
 DEFINITION sapiens genomic clone MCF7\_1-21P6, genomic survey sequence.  
 ACCESSION BZ611126  
 VERSION BZ611126  
 KEYWORDS GSS.  
 SOURCE BZ611126.1 GI:31519687  
 ORGANISM Homo sapiens (human)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 808)  
 AUTHORS Vollik,S., Zhao,S., Chin,K., Brebner,J.H., Herndon,D.R., Tao,Q.,  
 Kowbel,D., Huang,G., Lapuk,A., Kuo,W.-L., Magrane,G., de Jong,P.,  
 Gray,J.W. and Collins,C.  
 TITLE End-sequence profiling: Sequence-based analysis of aberrant genomes  
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 100 (13), 7696-7701 (2003)  
 MEDLINE 12798976  
 PUBMED 12798976  
 COMMENT Contact: Volik SV  
 Colin Collins, lab  
 UCSF Comprehensive Cancer Center

UCSF Box 0808, San Francisco, CA 94143-0808, USA  
 Tel: 415 502 7066  
 Fax: 415 502 5665  
 Email: svolik@cc.ucsf.edu  
 This clone is available from Amplicon Express  
 http://www.genomex.com  
 Class: BAC ends.

## FEATURES

Location/Qualifiers  
 1. .808  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /clone="MCF7\_1-21P6"  
 /sex="female"  
 /clone\_lib="Human MCF7 breast cancer cell line library  
 (MCF7\_1)"  
 /note="Vector: pCBAC1; Site\_1: HindIII; This library was  
 constructed from MCF7 breast\_cancer cell line by Amplicon  
 Express (http://www.genomex.com) using their standard  
 Procedure."

## ORIGIN

Query Match 8.0%; Score 240; DB 28; Length 808;  
 Best Local Similarity 73.6%; Pred. No. 7e-43;  
 Matches 306; Conservative 0; Mismatches 110; Indels 0; Gaps 0;  
 QY 298 ATACAGTGCCTATCAGCCGGGTGCAGTGGCTAGTGCCTGAAATCCCAACACTTTGGGAG 357  
 |||||  
 Db 210 AGAAATCTACTCTGGGCCGAGCACAGTGCCTCATGCTGTATCTTAGCACCTTTGGGAG 269  
 QY 358 GCGAGGTGGGTGCATCATTGAGGTTCAGAGTTCGAGACAGCTGGCCAAACACTGTGTA 417  
 |||||  
 Db 270 GCCAAGCGAGGTAGATCATTGAGGTTCAGAGTTCAGAGTTCAGAGTTCAGAGTTCAG 329  
 QY 418 AACCTGTCTCTACTAAAAATATAAAATTTAGCTGGGCATGGTGGTGCCTACCTGTAATC 477  
 |||||  
 Db 330 AACCTGTCTCTACTAAAAATATAAAATTTAGCTGGTGGTGGTGGTGGTGGTGGTGGTGG 389  
 QY 478 CGAGTACTTGGAGGTTCAGGCGAGAGATTCGCTTGAACCCGGGAGGAGATGTTGCGAG 537  
 |||||  
 Db 390 CCAGTACTTGGAGGTTCAGGCGAGAGATTCATTGAACCCAGGAGCGGAGGTTCGAG 449  
 QY 538 TGAACCAAGACTTGCACCTCGACTCGCTCTGGGCAACAGAGTGAAGCTTCCATCTCAAA 597  
 |||||  
 Db 450 TGAGCCGAGATTCCACCCTGTACTCCAGCTGGCAACAGAGTGAAGCTTCTCTCAAA 509  
 QY 598 CAAACAAACAAAGCAGTGCCTCATCTAGTATTGAGTATTGAGTGAAGTGAAGTGAAGC 657  
 |||||  
 Db 510 AAAAAAAAAAAG 569  
 QY 658 TTGTGCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 713  
 |||||  
 Db 570 TTTTTCATCTTGCAACAATCAATCACCCTGCTGTACCAACCAATCCACGTTCA 625

## RESULT 5

BZ599987 833 bp DNA linear GSS 08-JUN-2003  
 LOCUS WHACU62TR Human MCF7 breast cancer cell line library (MCF7\_1) Homo  
 DEFINITION sapiens genomic clone MCF7\_1-18K3, genomic survey sequence.  
 ACCESSION BZ599987  
 VERSION BZ599987.1 GI:31508449  
 KEYWORDS GSS.  
 SOURCE BZ599987.1  
 ORGANISM Homo sapiens (human)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 833)  
 AUTHORS Vollik,S., Zhao,S., Chin,K., Brebner,J.H., Herndon,D.R., Tao,Q.,  
 Kowbel,D., Huang,G., Lapuk,A., Kuo,W.-L., Magrane,G., de Jong,P.,  
 Gray,J.W. and Collins,C.  
 TITLE End-sequence profiling: Sequence-based analysis of aberrant genomes  
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 100 (13), 7696-7701 (2003)



```

DEFINITION K-EST0032169 S11SNU1 Homo sapiens cDNA clone S11SNU1-11-A07 5',
            mRNA sequence.
ACCESSION  BM754642
VERSION     BM754642.1  GI:19084260
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 551)
AUTHORS     Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
            Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
            Kim,Y.S.
TITLE       21C Frontier Korean EST Project 2001
JOURNAL     Unpublished (2002)
COMMENT     Contact: Kim YS
            Genome Research Center
            Korea Research Institute of Bioscience & Biotechnology
            52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
            Tel: +82-42-860-4470
            Fax: +82-42-860-4409
            Email: yongsung@mail.kribb.re.kr
            Plate: 11 row: A column: 07
            High quality sequence stop: 551.
            Location/Qualifiers
                1..551
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="S11SNU1-11-A07"
                /sex="M"
                /tissue_type="Stomach"
                /cell_type="Lymphoblast-like"
                /lab_host="SNU-1"
                /clone_lib="S11SNU1"
                /note="Organ: Stomach; Vector: pME18-FL3; Site 1: XhoI;
                Site 2: XhoI; The poly (A)+ RNA was dephosphorylated with
                bacterial alkaline phosphatase (BAP) and then decapped
                with tabacco acid pyrophosphatase (TAP). The decapped
                intact mRNA was ligated with DNA-RNA linker including SfiI
                site by treatment of T4 RNA ligase and the first strand
                cDNA was synthesized with Superscript II using SfiI
                oligo-dT primer. After first strand synthesis RNA was
                degraded by NaOH treatment and cDNA was amplified by PCR
                reaction. The PCR products were digested with SfiI and
                cloned into DraIII- digested pME18S-FL3 vector. The
                obtained cDNA vectors were used for transformation of
                competent cells E. coli Top10F' by electroporation method.
                The cDNA libraries constructed by this method are
                full-length enriched cDNA library."

FEATURES             source
     source
     1..657
     /organism="Homo sapiens"
     /mol_type="mRNA"
     /db_xref="taxon:9606"
     /tissue_type="CNCAP(3)T-225 cell line"
     /lab_host="DH10B (TI phage resistant)"
     /clone_lib="NIH_MGC_216"
     /note="Organ: Prostate; Vector: pT7T3 Pac; Site 1: EcoR I;
     Site 2: Not I; The library was constructed according
     Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
     1996. Denatured RNA was size fractionated on a 1% agarose
     gel. First strand cDNA synthesis was primed with oligo-dT
     primer containing a Not I site. Double strand cDNA was
     size selected according to mRNA size fraction, ligated
     with EcoR I adaptor, digested with Not I and then cloned
     directionally into pT7T3 Pac vector. The library tag
     sequence located between the Not I site and the polyA tail
     is TAGA. Tissue was provided by Tim Ratliff."

ORIGIN
Query Match      7.9%; Score 237.4; DB 12; Length 551;
Best Local Similarity 87.8%; Pred. No. 2.5e-42;
Matches 259; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

QY 314 GCCGGGTGAGTCGGCTAGTCCCTGAATCCACACTTTGGGAGCGGAGGTGGTGGAT 373
Db 209 GTCCGGCGGGTGGCTTATACCTGTATATCCAGCACTTTGGGAGGTGGGCGGAT 268

QY 374 CACTTGAGGTCCAGAGTTCGAGACCAAGCTGGCCCAACATGTCGAAACCTGTCTACTA 433
Db 269 CATCTGAGGTGGAGTTCGAGACCAAGCTTCGAAACATGTCGAAACCTGTCTACTA 328

QY 434 AATAATATAAATAGTCGGCAATGGTGGTGCCTGTGTAATCCAGATCTTGGGAGG 493
Db 329 AATAATATAAATAGTCGGCAATGGTGGTGCCTGTGTAATCCAGATCTTGGGAGG 388

QY 494 TTGAGCAGAGAGATCGCTTGAACCCGGGAGGAGATGTTGCAGTGAAACCAAGACTGTGC 553
Db 389 CTGAGGAGAGAGATCGCTTGAACCCGGGAGGAGATGTTGCAGTGAAACCAAGACTGTGC 448

```

```

QY 554 CACTGCACCTCCAGTCTGGGCAACAGAGTGAGCTCATCTCAAAACAACAA 608
Db 449 CACTGCACCTCCAGTCTGGGCGAGAGTGAGCTCATCTCAAAACAAAAA 503

RESULT 8
CF123536/c
LOCUS     CF123536             657 bp      mRNA      linear      EST 09-SEP-2003
DEFINITION UI-HF-CH0.2-aul-b-11-0-UI.r1 NIH_MGC_216 Homo sapiens cDNA 5', mRNA
sequence.
ACCESSION  CF123536
VERSION     CF123536.1  GI:33197844
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 657)
AUTHORS     Ronaldo,M.F., Lennon,G. and Soares,M.B.
TITLE       Normalization and subtraction: two approaches to facilitate gene
            discovery
JOURNAL     Genome Res. 6 (9), 791-806 (1996)
MEDLINE     97044477
PUBMED      889548
COMMENT     Contact: Soares, MB
            Coordinated Laboratory for Computational Genomics
            University of Iowa
            375 Newton Road , 4156 MEBRF, Iowa City, IA 52242, USA
            Tel: 319 335 8250
            Fax: 319 335 9565
            Email: bento-soares@uiowa.edu
            Tissue Procurement: Tim Ratliff
            cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
            cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
            DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
            Clone Distribution: Distribution information can be found at
            http://genome.uiowa.edu/distribution/humanfl.html
            The following repetitive elements were found in this cDNA
            sequence: 139-428, >ALU (matched complement) 344-450, >ALU (matched
            complement)
            Seq primer: pYX-5.
            Location/Qualifiers
                1..657
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /tissue_type="CNCAP(3)T-225 cell line"
                /lab_host="DH10B (TI phage resistant)"
                /clone_lib="NIH_MGC_216"
                /note="Organ: Prostate; Vector: pT7T3 Pac; Site 1: EcoR I;
                Site 2: Not I; The library was constructed according
                Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
                1996. Denatured RNA was size fractionated on a 1% agarose
                gel. First strand cDNA synthesis was primed with oligo-dT
                primer containing a Not I site. Double strand cDNA was
                size selected according to mRNA size fraction, ligated
                with EcoR I adaptor, digested with Not I and then cloned
                directionally into pT7T3 Pac vector. The library tag
                sequence located between the Not I site and the polyA tail
                is TAGA. Tissue was provided by Tim Ratliff."

ORIGIN
Query Match      7.9%; Score 237.4; DB 14; Length 657;
Best Local Similarity 86.5%; Pred. No. 2.6e-42;
Matches 262; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 314 GCCGGGTGAGTCGGCTAGTCCCTGAATCCACACTTTGGGAGCGGAGGTGGTGGAT 373
Db 427 GCCGGCGGAGTGGCTTATATCCAGCACTTTGGGAGCGGAGGTGGGCGAGC 368

QY 374 CACTTGAGGTCCAGAGTTCGAGACCAAGCTGGCCCAACATGTCGAAACCTGTCTACTA 433
Db 367 CATCTGAGTCCAGAGTTCGAGACCAAGCTGGCCCAACATGTCGAAACCTGTCTACTA 308

```



ORIGIN									
	Query Match	7.9%	Score 237;	DB 28;	Length 690;				
	Best Local Similarity	84.2%;	Pred. No. 3.2e-42;						
	Matches 267;	Conservative 0;	Mismatches 50;	Indels 0;	Gaps 0;				
	/cell_type="Lymphocytes" /clone_lib="RPCI-11" /note="Vector: pBACe3.6; Site 1: EcoRI; Site 2: EcoRI; RPCI11 Human Male BAC Library"								
Qy	295	GTGATAACAGTGCCTCAGCGGGTGCAGTGGCTAGTGCCTGAAATCCCAACACTTTGG	354						
Db	21	GTTTTAAAGACATATCAGCAGCGGGTGGCTTATGCTGTAATCCGACGACTTTGG	80						
Qy	355	GAGCGGAGGTGGGTGGATCACTTGAGGTCAGGAGTTCGAGACGAGCCTGGCCAACTGG	414						
Db	81	GAGGCTCAGGAAAGTGGATCATCAGGTGAGGTCAGGATTTCAAGACCAGCCTGGACAACTGG	140						
Qy	415	TGAAACCCCTGTCTCTACTTAAAAATATATAAAATTAGCTGGGCATCGTGTGGTACCTGTGA	474						
Db	141	CAAAACCCCTGTCTCTACTTAAAAATACAAAAATTAGCCGGAGTGGTGTGACACTGTGA	200						
Qy	475	ATCCAGATACTTGGGAGGTTGAGGCGAGGAGAAATCGCTTGAAACCCGGAGGACAGATTTG	534						
Db	201	ATCCAGCTACTTCGGGAGGCTGAGCGAGGAGAAATCGCTTGAACTGGGAGGACAGGTTG	260						
Qy	535	CAGTGAACCAAGACTGTGCCACTGCACCTCCAGTCTGGGCAACAGAGTGAGCCTCCATCTC	594						
Db	261	CAGTGAGCCAAGATCGTACCACTCCAGCTCCAGCCTGGGCGACAGAGTGAACCTCTGCTC	320						
Qy	595	AAACAAACAAACAAAA	611						
Db	321	AAACAAACAAACAAAA	337						

RESULT 11  
AW471332  
LOCUS  
DEFINITION  
          494 bp          mRNA          linear          EST 24-FEB-2000  
          xw57hl1.x1 NCI\_CGAP\_Pan1 Homo sapiens cDNA clone IMAGE:2832165 3'  
          similar to contains Alu repetitive element;; mRNA sequence.  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
          AW471332.1 GI:7041438  
          EST.  
          Homo sapiens (human)  
          Homo sapiens  
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
          1 (bases 1 to 494)  
AUTHORS  
          NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
TITLE  
          National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
          Tumor Gene Index  
JOURNAL  
          Unpublished (1997)  
COMMENT  
          Contact: Robert Strausberg, Ph.D.  
          Email: cgapsb-r@mail.nih.gov  
          Life Technologies catalog #: 11548-013  
          DNA sequencing by: Washington University Genome Sequencing Center  
          Clone distribution: NCI-CGAP clone distribution information can be  
          found through the I.M.A.G.E. Consortium/LLNL at:  
          [www-bio.llnl.gov/bbrp/image/image.html](http://www-bio.llnl.gov/bbrp/image/image.html)  
          Seq primer: -400P from Gibco  
          High quality sequence stop: 417.

ORIGIN	Query Match	7.9%	Score 236.8;	DB 10;	Length 494;
	Best Local Similarity	84.7%	Fred. NO. 3.4e-42;	Mismatches 48;	Indels 0; Gaps 0;
	Matches 265;	Conservative 0;			
QY	315	CGGGTGCAGTGGCTAGTGCCTGAATCCAAACACTTTTGGGAGCGGAGGTGGGTGATC	374		
Db	181	CGGGTGTGGTGGCTCATGCTGTAATTCAGCACATTTGGGAGGCTGAAGTGGGAGATC	240		
QY	375	ACTTGAGGTTCAGGAGTTTCGAGACAGCCTGGCCAAACATGGTGAACCCCTGTCTCTACTAA	434		
Db	241	ACTTCAGGTTCAGGAGTTTCGAGACAGCCTGGCCAAATATGCTGAAACCCCGCTCTCCACTAA	300		
QY	435	AAATATAAAAATTAGCTGGGCATGGTGGTACCTGTGTAATCCGAGATACACTTGGGAGGT	494		
Db	301	AAATATAAAAATTAGCTGGGTGGTGGTAGAGTGTGCCCTGTAAATCCCACTACTTGGGAGGC	360		
QY	495	TGAGGCAGGAGAAATCGCTTCAACCCGGGAGGAGAGTGTTCAGTGAACCAAGACTGTGCC	554		
Db	361	TGAGGCAGGAGAAATCGCTCAACCTGGGAGGAGAGTGTTCAGTGAACCAAGACTGTGCC	420		
QY	555	ACTGCACCTCAGTCTGGGCAACAGAGTGAGCCCTCCATCTCAAAACAAACAAAAAGCA	614		
Db	421	ACTGCACCTCAGGCTGGGCAACAGAGTGAGCCCTCCATCTCAAAACAAACAAAAAGCA	480		
QY	615	GTGCCCATCATGT 627			
Db	481	CGTGCCTTCTTAT 493			
RESULT 12					
BQ880197/c					
LOCUS	BQ880197	932 bp	mRNA	linear	EST 16-AUG-2002
DEFINITION	AGENCYCOURT 8122269 lupski dorsal root ganglion Homo sapiens cDNA				
ACCESSION	BQ880197				
VERSION	BQ880197.1	GI:22272205			
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
TITLE	1 (bases 1 to 932)				
JOURNAL	NH-MGC <a href="http://mgc.nci.nih.gov/">http://mgc.nci.nih.gov/</a> .				
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)				
	Unpublished (1999)				
	Contact: Robert Strausberg, Ph.D.				
	Email: <a href="mailto:cgapbs-r@mail.nih.gov">cgapbs-r@mail.nih.gov</a>				
	Tissue Procurement: Dr. James R. Lupski				
	cDNA Library Preparation: Life Technologies, Inc.				
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)				
	DNA Sequencing by: Agencourt Bioscience Corporation				
	Clone distribution: MGC clone distribution information can be				
	found through the I.M.A.G.E. Consortium/LLNL at:				
	<a href="http://image.llnl.gov">http://image.llnl.gov</a>				
	Plate: LLMI3560 row: m column: 14				
	High quality sequence stop: 604.				
FEATURES					
	Location/Qualifiers				
	1..932				

```

/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6179509"
/sex="male"
/tissue_type="dorsal root ganglia"
/dev_stage="adult, 36 yr"
/lab_host="DH10B"
/clone_lib="lupski_dorsal_root_ganglion"
/notes="Vector: PCMW-SPORT6 (Life Technologies); Site 1:
NotI; Site 2: SalI; cDNA made by oligo-dT priming.
Directionally cloned using the following adaptors:

```

5'-TCGACCCACCGCTCCG-3' and  
5'-GACTAGTCTAGATCGCAGCGCGCCCT(15)-3'. Size selected >  
1 kb for average insert length 1.7 kb. This is a primary  
library, non-amplified. Library constructed by Life  
Technologies and donated by J. Lupeki, M.D./Ph.D. (Baylor  
College of Medicine) and is available through Life  
Technologies."

## ORIGIN

Query Match 7.9%; Score 236.8; DB 13; Length 932;  
Best Local Similarity 63.4%; Pred. No. 3.8e-42;  
Matches 378; Conservative 0; Mismatches 217; Indels 1; Gaps 1;  
QY 2407 GCTCAGCGTCCCATCTTAGCAGCTTTGGAGGCTGAGTGGTGTGATCTCAGTCAG 2466  
DB 691 GCTCAGCGCTATAATCCCGAGCACTTTGGAGGCTGAGGAGATGATCACCGTGGT 632  
QY 2467 GAGTTCAAGACGAGCTAGACAAACGTAGAGAAACCCCATCTCTACTGAAGATA-TAAAT 2525  
DB 631 GAGTTGAGACCATCTCGCCACATGCGAAACCCCGTCTCTACTGAATAACAAAT 572  
QY 2526 TAGCTGGCGTAGTGGCATATGCCGTGTAATCCAGCTAGTCGGGAAGCTGAGCGAGGAGA 2585  
DB 571 TAGCAGCGGTGGTGGCAGGCGCCCTGTAGTCCAGCTACTTTGGGAGGCTGAGCGAGGAGA 512  
QY 2586 ATCGCTTGAACCCGAGAGCGAGGTTGGGTGAGCTGAGATCGTGCATTCACCTCCAG 2645  
DB 511 ATCACTTGAACCCGGGAGGAGGTTGAGTGGAGCGGAGATCGTGCATTCACCTCCAG 452  
QY 2646 CTGGGTGACAAAGCAAGACCTCGTCTCAATAATAATAATACAAACAGAGGAGC 2705  
DB 451 CTGGGAGCAGAGGAGACTCGCTCCAGAAAAAAGAAAGAACTAATAAATATTG 392  
QY 2706 CTGGGTGATCCAGCTACTTCTTTCAGGAGATGTACTCCCTTACCCAGGCGAAGG 2765  
DB 391 GGAATACTTCTTGTACTAATACATACCTAAACCTTTAATACTAGAACAGATGCCTACATC 332  
QY 2766 ATGGGAGAACAGTTTGTATTGATTTATTGAGCAGCTACTGAGTCTCATCCCTGGGC 2825  
DB 331 CAGTATTAAATAGTAAATATCAACCAAGTATACCTTTTATTTTACCTACACTTTA 272  
QY 2826 TAGGTGGAATGGATCAGATGGAGCTCGAAGAGTCCCGCTCAGGGAACCTCAGTAGAAA 2885  
DB 271 CTAAATAGTTGAAAAAAGCGCAAGAGGAGGATCATCTCTCTCCAGTGATATAA 212  
QY 2886 GAAGAGGAATCGCGCGCGGCTGAGTCAAGCTGTAAATCCCAACACTTTGGAGGCT 2945  
DB 211 AGAAGATGATGGGCTGGCGGCTGAGCTCAAGCTGTAAATCCAGCACTTTGGGAAGCC 152  
QY 2946 GAGGTGGGTGATCACAAGGTCAGGAGATCGAGACCATCTCGGTAAACACAGTGAA 3001  
DB 151 AAGGCTGGCAGATCAGAGTCAGAGATCGAGACCATCTCGGCAACATGTTGAA 96

RESULT 13  
A0045514  
LOCUS A0045514 498 bp DNA linear GSS 14-APR-1999  
DEFINITION RPC111-3387.TJ RPCI-11 Homo sapiens genomic clone RPCI-11-3387,  
genomic survey sequence.  
ACCESSION A0045514  
VERSION A0045514.1 GI:3314441  
KEYWORDS GSS.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 498)  
AUTHORS Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,  
Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., de Jong,P. and  
Venter,J.C.  
Use of BAC End Sequences for Sequence-Ready Map Building (1998)  
JOURNAL Unpublished (1998)  
COMMENT Contact: Mark Adams

Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: mdadams@igr.org  
Clones are derived from the human BAC library RPCI-11. For BAC  
library availability, please contact Pieter de Jong  
(pieterdejong.med.buffalo.edu). Clones may be purchased from  
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from  
Research Genetics (info@resgen.com). BAC end search page:  
http://www.tigr.org/tdb/hungen/bac\_end\_search/bac\_end\_search.html  
Class: BAC ends.

## FEATURES

## source

Location/Qualifiers  
1..498  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="GDB:7512390"  
/db\_xref="taxon:9606"  
/clone="RPCI-11-3387"  
/sex="Male"  
/cell\_type="Lymphocytes"  
/clone\_lib="RPCI-11"  
/note="Vector: pBACe3.6; Site\_1: EcoRI; Site\_2: EcoRI;  
RPC111 Human Male BAC Library"

## ORIGIN

Query Match 7.9%; Score 236; DB 28; Length 498;  
Best Local Similarity 84.2%; Pred. No. 5.1e-42;  
Matches 266; Conservative 0; Mismatches 50; Indels 0; Gaps 0;  
QY 236 TGATAACAGTGCCCATCAGCGGGTGAGTGGCTAGTCCCTGAAATCCCAACACTTTGGG 355  
DB 137 TTATGATAGTATTCCTCAGCTGGGCGGCTGGGTACGCTGTAATCCAGCAGCTTTGGG 196  
QY 356 AGCGGAGGTGGGTGGATCAGTTCAGGTTCAGAGTTCGAGACCAAGCTGGGCCAACATGCT 415  
DB 197 AGGCCAAGTCAGTGTGATCAGTTCAGGTTCAGAGTTCGAGACCAAGCTGGGCCAACATGCT 256  
QY 416 GAAACCTCTCTACTAAAAATATAAAAAATAGCTGGGCATGGTGGCTGAGTACCTGTAA 475  
DB 257 GAAACCTCTTTGTAATAAATAACAAAAATTAAGTGGGCATGGTGGGCACCTATAA 316  
QY 476 TCCAGATACCTGGAGGTTCAGGAGGAGTTCGTTGAACCCGGGAGGAGTGTTCG 535  
DB 317 TCCAGCTACTTTGGAGGCTGAGGAGAGATCACTTGAACCTGGGAGGAGTGTTCG 376  
QY 536 AGTGAACCAAGACTGTGCCACTGCACTCCAGTTCGGCAACAGAGTGGCTCCATCTCA 595  
DB 377 AGTGAGCGAGATCGCGCATTTGCACTCCAGCTAGGAGCAAGTGAAGTCCATCTCA 436  
QY 596 AACAAACAAACAAAAA 611  
DB 437 AAAAAAAAAAAAAAAAAA 452

## RESULT 14

## A0035209

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

AG035209 656 bp DNA linear GSS 01-NOV-2001  
Pan troglodytes DNA, clone: PTB-010M15.F, genomic survey sequence.  
AG035209  
AG035209.1 GI:16562082  
GSS  
Pan troglodytes (chimpanzee)  
Pan troglodytes  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.  
Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,  
Tokui,Y., Watanabe,H. and Sakaki,Y.  
BAC end sequences of Library PTB  
Unpublished  
2 (bases 1 to 656)



**AUTHORS** Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,  
**TITLE** Totoki, Y., Watanabe, H. and Sakaki, Y.  
**JOURNAL** Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical  
 and Chemical Research (RIKEN), Genomic Sciences Center (GSC);  
 1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
 (E-mail: chimbes@gsc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/  
 Tel: 81-45-503-9111, Fax: 81-45-503-9170)  
**COMMENT** Clones are derived from the chimpanzee BAC library PTB This BAC end  
 was generated during the R&D process and may have higher chance of  
 clone tracking errors.  
**PRIMERS**  
 Sequencing: -21M13  
**LIBRARY**  
 Vector : pKS145  
 R.Site 1 : SacI  
 R.Site 2 : SacI.  
**FEATURES**  
 source Location/Qualifiers  
 1. .656  
 /organism="Pan troglodytes"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9598"  
 /clone="PTB-010M15.F"  
 /sex="male"  
 /cell\_type="lymphoblast"  
 /clone\_lib="PTB Chimpanzee Male BAC Library"  
**ORIGIN**  
 Query Match 7.9%; Score 236; DB 29; Length 656;  
 Best Local Similarity 86.7%; Pred. No. 5.3e-42;  
 Matches 260; Conservative 0; Mismatches 40; Indels 0; Gaps 0;  
 QY 313 AGCCGGGTGCAGTGGCTAGTCCGTGAATCCACACTTTGGAGGCGGAGGTGGGTGGA 372  
 Db 193 AGCCAGCGCAGTAGCTTCATGCCGGTAAATCCAGCACTTTGGAGGCGGAGGTGGGA 252  
 QY 373 TCACCTTGGAGTCAGGAGTTCAGACACAGCTGGCCACATGGTGMAACCTGTCTCTACT 432  
 Db 253 TCACCTGAGTTCAGGAGTTCAGACACAGCTGGCCACATGGTGMAACTGTCTCTACT 312  
 QY 433 AAAAAATATAAAATTAGCTGGGATGGTGGTGGTACCTGTAAATCCAGATATCTTGGGAG 492  
 Db 313 AAAAAATACAAAAATTAGCTGGGATGGTGGTGGTACCTGTAAATCCAGTACTAGGAAG 372  
 QY 493 GTTGAGCGGAGCAATCGTTGACCCGAGCAGCAGATGTGCAAGTGAACCAAGACTGTG 552  
 Db 373 GCTGAGCGGAGGAATCGTTGAAACCGGAGCAGAGTTGCAAGTGCAGCGGAGTCGTG 432  
 QY 553 CCACCTGCATCCCACTCTGGGCAACAGAGTCAGCTCCATCTCAAAACAAACAAACAAAG 612  
 Db 433 CCACCTGTACTCCCACTGGGCGACTGAGTGAGATCCATCTCAGTCATCAACCAAGTCAG 492  
**RESULT 15**  
**LOCUS** AQ590440  
**DEFINITION** HS\_5388\_A1\_D12\_T7A RPCI-11 Human Male BAC Library Homo sapiens  
 genomic clone Plate=964 Col=23 Row=G, genomic survey sequence.  
**ACCESSION** AQ590440  
**VERSION** AQ590440.1 GI:5022092  
**KEYWORDS** GSS.  
**SOURCE** Homo sapiens (human)  
**ORGANISM**  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
**REFERENCE**  
**AUTHORS** Mahairas, G.G., Wallace, J.C., Smith, K., Swartzell, S., Holzman, T.,  
 Keller, A., Shaker, R., Furlong, J., Young, J., Zhao, S., Adams, M.D. and  
 Hood, L.  
**TITLE** Sequence-tagged connectors: A sequence approach to mapping and  
 scanning the human genome  
**JOURNAL** Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)  
**MEDLINE** 99380589

**PUBMED** 10449764  
**COMMENT**

Contact: Mahairas GG, Wallace JC, Hood L  
 High Throughput Sequencing Center  
 University of Washington  
 401 Queen Anne Avenue North, Seattle, WA 98109, USA  
 Tel: (206) 616-3618  
 Fax: (206) 616-3887  
 Email: jwallaceu.washington.edu  
 Clones are derived from the human BAC library RPCI-11. For BAC  
 library availability, please contact Pieter de Jong  
 (pieter@dejong.med.buffalo.edu). Clones may be purchased from  
 BACPAC Resources (http://bacpac.med.buffalo.edu/ordering\_bac.htm)  
 or from Resear h Genetics (info@resgen.com). BAC end Web Server:  
 http://www.husc.washington.edu  
 Plate: 964 row: G column: 23  
 Seq primer: T7  
 Class: BAC ends  
 High quality sequence stop: 466.

**FEATURES**  
 source Location/Qualifiers  
 1. .466

/organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /clone="Plate=964 Col=23 Row=G"  
 /sex="male"  
 /clone\_lib="RPCI-11 Human Male BAC Library"  
 /notes="Vector: pBACE3.6; Site 1: EcoRI; Site 2: EcoRI;  
 Male blood DNA was isolated from one randomly chosen donor  
 and partially digested with a combination of EcoRI and  
 EcoRI Methylase. Size selected DNA was cloned into the  
 pBACE3.6 vector at EcoRI sites"

#### ORIGIN

Query Match 7.9%; Score 235.8; DB 28; Length 466;  
 Best Local Similarity 87.5%; Pred. No. 5.6e-42;  
 Matches 258; Conservative 0; Mismatches 37; Indels 0; Gaps 0;  
 QY 315 CCGGGTGCAGTGGCTAGTGCCTGAAATCCCAACACTTTGGAGGCGGAGGTGGGTGATC 374  
 Db 61 CCAGGCGCGTGGCTCACGCTGTAATCCAGCACCTTTGGGAGGCATAGTGGGTGATC 120  
 QY 375 ACTTGAGGTCAGGAGTTCGAGACGAGCTGGCCAAACATGCTGAACCCCTCTCTACTAA 434  
 Db 121 ACTTGAGGTCAGGAGTTCGAGACGAGCTGGCCAAATATGCTGAACCCCTCTGTACTAA 180  
 QY 435 AAATATAAAATTAGCTGGGCGCATGGTGGTGGTACCTGTAAATCCAGATATCTTGGGAGGT 494  
 Db 181 AAATATAAAATTAGCCAGGCTTGGCGGACGCTGTAGTCCAGCTACTTGGGAGGC 240  
 QY 495 TGGGGCAGGAGATCGCTTGAACCCGGGAGGAGAGATGTTGCAAGTGAACCAAGACTGTGCC 554  
 Db 241 TGAGGCGAGGAGATCGCTTGAACCCAGGAGGAGAGGTTGCAAGTGAAGATTGTGCC 300  
 QY 555 ACTGCACCTCCAGTCTGGGCAACAGAGTGAGCTCCATCTCAACCAACCAACAA 609  
 Db 301 ACTGCACCTCCAGCTAGGCAACAGAGTGAGACTCTCTCTCAAAAAA 355

Search completed: September 17, 2004, 18:58:24  
 Job time : 7317.95 secs

***This Page Blank (uspto)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2004, 09:14:41; Search time 121 Seconds  
(without alignments)  
1314.225 Million cell updates/sec

Title: US-10-669-693-2

Perfect score: 2615

Sequence: 1 MEATGTWALLALALLLT.....PPAFQLQVRPTDLHSTTQTR 504

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- SPTREMBL\_25.\*
- 1: sp\_archaea.\*
  - 2: sp\_bacteria.\*
  - 3: sp\_fungi.\*
  - 4: sp\_human.\*
  - 5: sp\_invertebrate.\*
  - 6: sp\_mammal.\*
  - 7: sp\_mhc.\*
  - 8: sp\_organelle.\*
  - 9: sp\_phage.\*
  - 10: sp\_plant.\*
  - 11: sp\_rodent.\*
  - 12: sp\_virus.\*
  - 13: sp\_vertebrate.\*
  - 14: sp\_unclassified.\*
  - 15: sp\_rvirus.\*
  - 16: sp\_bacteriap.\*
  - 17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	%	Length	DB ID	Description
1	2615	100.0	504	4	Q9BZ66	Q9bZ66 homo sapien
2	2527	96.6	564	4	Q96SQ9	Q96sq9 homo sapien
3	2060.5	78.8	501	11	Q9DBX6	Q9dbx6 mus musculus
4	1313	50.2	491	6	Q29532	Q29532 oryctolagus
5	1301	49.8	493	6	Q8SQ67	Q8sq67 sus scrofa
6	1300	49.7	323	11	Q8Q027	Q8q027 mus musculus
7	1298	49.6	491	6	Q29516	Q29516 oryctolagus
8	1296	49.6	491	11	Q9WDD0	Q9wd0 mus musculus
9	1261.5	48.2	494	11	Q9QZ50	Q9qz50 mesocricetu
10	1257	48.1	491	11	Q64584	Q64584 rattus norv
11	1255	48.0	491	11	Q8QXG6	Q8qxg6 mesocricetu
12	1247	47.7	491	11	Q64463	Q64463 mus musculus
13	1243.5	47.6	494	11	Q9WV19	Q9wv19 mus musculus
14	1242.5	47.5	494	6	Q8SQ68	Q8sq68 sus scrofa
15	1236	47.3	493	11	Q8QXK1	Q8qxk1 mus musculus
16	1236	47.3	494	11	Q91X75	Q91x75 mus musculus

17	1233	47.2	494	11	Q91XG2	Q91xg2 mus musculus
18	1226	46.9	493	11	Q9R0R6	Q9r0r6 cricetus
19	1224	46.8	494	11	Q9JJ02	Q9jj02 rattus norv
20	1221.5	46.7	494	11	Q9Z1H8	Q9z1h8 cricetus
21	1214.5	46.4	495	11	Q64583	Q64583 rattus norv
22	1210	46.3	491	11	Q64460	Q64460 mus musculus
23	1210	46.3	495	11	Q9JM91	Q9jm91 cricetus
24	1199.5	45.9	488	6	Q7YSF7	Q7ysf7 sus scrofa
25	1180.5	45.1	494	13	Q92129	Q92129 xenopus lae
26	1175.5	45.0	494	4	Q13120	Q13120 homo sapien
27	1168	44.7	494	11	Q8QZM4	Q8qzm4 mus musculus
28	1166.5	44.6	492	11	Q8VCM9	Q8vcw9 mus musculus
29	1164	44.5	487	11	Q8CIE7	Q8ciel gorilla gor
30	1156	44.2	491	6	Q8WNE1	Q8wnel gorilla gor
31	1152.5	44.1	494	6	Q8SQ66	Q8sq66 sus scrofa
32	1127	43.1	494	13	Q8QFT4	Q8qft4 gallus gall
33	1125	43.0	495	4	Q8WJ2	Q8wvj2 homo sapien
34	1089.5	41.7	490	6	Q8SQ65	Q8sq65 sus scrofa
35	1085.5	41.5	487	4	Q16756	Q16756 homo sapien
36	1085	41.5	496	13	Q7ZX81	Q7zx81 xenopus lae
37	1083.5	41.4	490	11	Q8C7K2	Q8c7k2 mus musculus
38	1074.5	41.1	490	11	Q91WN9	Q91wn9 mus musculus
39	1071.5	41.0	477	4	Q16872	Q16872 homo sapien
40	1069	40.9	474	11	Q9CVC8	Q9cvc8 mus musculus
41	1056.5	40.4	490	11	Q9BDD9	Q9bdb9 mus musculus
42	1052.5	40.2	485	4	Q9UC29	Q9ucz9 homo sapien
43	1051.5	40.2	490	4	Q8WBI	Q8wbi homo sapien
44	1046.5	40.0	495	11	Q91Y29	Q91y29 rattus norv
45	1043.5	39.9	490	11	Q91X77	Q91x77 mus musculus

#### ALIGNMENTS

#### RESULT 1

Q9BZ66 ID Q9BZ66 PRELIMINARY; PRT; 504 AA.  
AC Q9BZ66;  
DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DB Cytochrome P450 2S1 (Similar to cytochrome P450, subfamily IIS, polypeptide 1).  
DE Polypeptide 1).  
GN CP2S1.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21092856; PubMed=11181079;  
RA Rylander T., Neve E., Ingelman-Sundberg M., Oscarson M.;  
RT "Identification and Tissue Distribution of the Novel Human Cytochrome P450 2S1 (CYP2S1)";  
RL Biochem. Biophys. Res. Commun. 281:529-535(2001).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Blood;  
RA Strausberg R.;  
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.  
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.  
DR EMBL; AF335278; AAK13498.1; -;  
DR EMBL; BC033691; AAH33691.1; -;  
DR PIR; JC7613; JC7613.  
DR HSSP; P00179; 1DT6.  
DR GO; GO:0005783; C:endooplasmic reticulum; NAS.  
DR GO; GO:0004497; F:monooxygenase activity; NAS.  
DR GO; GO:0006118; P:electron transport; NAS.  
DR InterPro; IPR001128; Cytochrome\_P450.  
DR Pfam; PF00067; P450; 1.  
DR PRINTS; PR00385; P450.  
DR PROSITE; PS00086; CYTOCHROME P450; 1.  
KW Heme; Monooxygenase; Oxidoreductase.

```
KW Hypothetical protein; Heme; Monooxygenase; Oxidoreductase.
SQ SEQUENCE 564 AA; 62231 MW; 03B2E9376A494D12 CRC64;

Query Match 96.6%; Score 2527; DB 4; Length 564;
Best Local Similarity 100.0%; Pred. No. 1.2e-192;
Matches 487; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEATGTWALLALALLLLLTALSSTRARGHLPPGPTPLPLGNLLQLRPGALYSGLMRL 60
DB 1 MEATGTWALLALALLLLLTALSSTRARGHLPPGPTPLPLGNLLQLRPGALYSGLMRL 60

QY 61 SKKYGPFVTIYIGPWRPVVVLVGQAVREALGGQAEFSGRTVAMLEGTDFGHGVFFSN 120
DB 61 SKKYGPFVTIYIGPWRPVVVLVGQAVREALGGQAEFSGRTVAMLEGTDFGHGVFFSN 120

QY 121 GERWQLRKFTMLALRDLMGKREGEELIQAEARCLVETFOGTEGRPPDPSLLLAQATSN 180
DB 121 GERWQLRKFTMLALRDLMGKREGEELIQAEARCLVETFOGTEGRPPDPSLLLAQATSN 180

QY 181 VVCSLLFGLRFSYEDKEFOAVVRAAGGTLGVSSQGGQTYEMFSWFLRPLPGPHKQLLHH 240
DB 181 VVCSLLFGLRFSYEDKEFOAVVRAAGGTLGVSSQGGQTYEMFSWFLRPLPGPHKQLLHH 240

QY 241 VSTLAFTVRQVQOQHGNDASGPARDLVDAFLKMAQEEQNPTEFTNKMLMTVIYLL 300
DB 241 VSTLAFTVRQVQOQHGNDASGPARDLVDAFLKMAQEEQNPTEFTNKMLMTVIYLL 300

QY 301 FAGTMTVSTTVGYTLTLLMKYPHVQKWREELNRELGAGQAPSLGDRTRLPTDVLHEA 360
DB 301 FAGTMTVSTTVGYTLTLLMKYPHVQKWREELNRELGAGQAPSLGDRTRLPTDVLHEA 360

QY 361 QRLALLVPNGIPTLMTTRFRGYTLPGQTEVPFLGLSILHDPNIFKHPEEFNPDRLDA 420
DB 361 QRLALLVPNGIPTLMTTRFRGYTLPGQTEVPFLGLSILHDPNIFKHPEEFNPDRLDA 420

QY 421 DGRFRKHEAFLPFLSKRVCGLGELAKAELFFFTTILQAFSLSPCPDPTLSLKPTVSG 480
DB 421 DGRFRKHEAFLPFLSKRVCGLGELAKAELFFFTTILQAFSLSPCPDPTLSLKPTVSG 480

QY 481 LFNIPPA 487
DB 481 LFNIPPA 487

RESULT 3
Q9DBX6 PRELIMINARY; PRT; 501 AA.
AC Q9DBX6;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE 1200011CL5RIK protein (Similar to cytochrome P450 2B1).
GN CYP2B1 OR 1200011CL5RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=C57BL/6J; TISSUE=Lung;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadoya K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
```

```
SQ SEQUENCE 504 AA; 55816 MW; 853513370F65E25A CRC64;

Query Match 100.0%; Score 2615; DB 4; Length 504;
Best Local Similarity 100.0%; Pred. No. 1e-199;
Matches 504; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEATGTWALLALALLLLLTALSSTRARGHLPPGPTPLPLGNLLQLRPGALYSGLMRL 60
DB 1 MEATGTWALLALALLLLLTALSSTRARGHLPPGPTPLPLGNLLQLRPGALYSGLMRL 60

QY 61 SKKYGPFVTIYIGPWRPVVVLVGQAVREALGGQAEFSGRTVAMLEGTDFGHGVFFSN 120
DB 61 SKKYGPFVTIYIGPWRPVVVLVGQAVREALGGQAEFSGRTVAMLEGTDFGHGVFFSN 120

QY 121 GERWQLRKFTMLALRDLMGKREGEELIQAEARCLVETFOGTEGRPPDPSLLLAQATSN 180
DB 121 GERWQLRKFTMLALRDLMGKREGEELIQAEARCLVETFOGTEGRPPDPSLLLAQATSN 180

QY 181 VVCSLLFGLRFSYEDKEFOAVVRAAGGTLGVSSQGGQTYEMFSWFLRPLPGPHKQLLHH 240
DB 181 VVCSLLFGLRFSYEDKEFOAVVRAAGGTLGVSSQGGQTYEMFSWFLRPLPGPHKQLLHH 240

QY 241 VSTLAFTVRQVQOQHGNDASGPARDLVDAFLKMAQEEQNPTEFTNKMLMTVIYLL 300
DB 241 VSTLAFTVRQVQOQHGNDASGPARDLVDAFLKMAQEEQNPTEFTNKMLMTVIYLL 300

QY 301 FAGTMTVSTTVGYTLTLLMKYPHVQKWREELNRELGAGQAPSLGDRTRLPTDVLHEA 360
DB 301 FAGTMTVSTTVGYTLTLLMKYPHVQKWREELNRELGAGQAPSLGDRTRLPTDVLHEA 360

QY 361 QRLALLVPNGIPTLMTTRFRGYTLPGQTEVPFLGLSILHDPNIFKHPEEFNPDRLDA 420
DB 361 QRLALLVPNGIPTLMTTRFRGYTLPGQTEVPFLGLSILHDPNIFKHPEEFNPDRLDA 420

QY 421 DGRFRKHEAFLPFLSKRVCGLGELAKAELFFFTTILQAFSLSPCPDPTLSLKPTVSG 480
DB 421 DGRFRKHEAFLPFLSKRVCGLGELAKAELFFFTTILQAFSLSPCPDPTLSLKPTVSG 480

QY 481 LFNIPPAQLQVRPTDLHSTTQTR 504
DB 481 LFNIPPAQLQVRPTDLHSTTQTR 504

RESULT 2
Q96SQ9 PRELIMINARY; PRT; 564 AA.
AC Q96SQ9;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein FLJ14699.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Ishibashi T., Fujimori K.,
RA Tanai H., Kimata M., Watanabe M., Hiraoka S., Ishii S., Kawai Y.,
RA Saito K., Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K.,
RA Masuko Y., Kanehori K.;
RA "NEDO human cDNA sequencing project.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR EMBL: AK027605; BAB55227.1; -.
DR GenBank: U15654; CYP2S1.
DR GO: 0004497; F:monooxygenase activity; IEA.
DR GO: 0006118; P:electron transport; IEA.
DR InterPro: IPR001128; Cytochrome_P450.
DR Pfam: PF00067; p450; 1.
DR PRINTS: PR00385; P450.
DR PROSITE: PS00086; CYTOCHROME_P450; 1.
```

RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seva T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,  
 RA Wynshaw-Boris A., Yoshida K., Haegawa Y., Kawaji H., Kohsaki S.,  
 RT Hayashizaki Y.,  
 RL "Functional annotation of a full-length mouse cDNA collection."; [2]  
 RN Nature 409:685-690(2001).  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Lung;  
 RX MEDLINE=22354683; PubMed=12466851;  
 RA The FANTOM Consortium,  
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 RT 60,770 full-length cDNAs";  
 RL Nature 420:563-573(2002).  
 CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.  
 DR EMBL; AK004699; BAB33484.1; -;  
 DR EMBL; AK087069; BAC39794.1; -;  
 DR HSSP; P00179; 1DT6.  
 DR MGD; MGI:1921384; Cyp2a1.  
 DR GO; GO:0004497; F:monooxygenase activity; IEA.  
 DR GO; GO:0006118; P:electron transport; IEA.  
 DR InterPro; IPR001128; Cytochrome\_P450.  
 DR Pfam; PF00067; P450; 1.  
 DR PRINTS; PR00385; P450.  
 KW Heme; Monooxygenase; Oxidoreductase.  
 SQ SEQUENCE 501 AA; 55631 MW; D12575F3F8D7B019 CRC64;

Query Match 78.8%; Score 2060.5; DB 11; Length 501;  
 Best Local Similarity 78.8%; Pred. No. 1.4e-155;  
 Matches 391; Conservative 50; Mismatches 54; Indels 1; Gaps 1;

QY 1 MEATGTTALLLA-LALLLLLTALSGTRAGHLPPGPTPLPLGNLLQRLGALYGLMR 59  
 Db 1 MEASTWALLALLLLLSLTFTPTARGYLPFGPTPLPLGNLLQRLGALYGLLR 60  
 QY 60 LSKYGVFTYILGPWPVPVVLVQEAAREALGQAEFEFGRTVAMLEGTFDGHHGVFFS 119  
 Db 61 LSKYGVFTYILGPWPVPVVLVQEAAREALGQAEFEFGRTVAMLEGTFDGHHGVFFA 120  
 QY 120 NGERWRLKFTMLALDLGKRGEGELIQAEARCLVETFGTGRFPDPSLLAQAATS 179  
 Db 121 NGERWKQLRKFTLLALDLGKRGEGELIQAEVQSLVEAFQKTEGFPNPSMLAQAATS 180  
 QY 180 NVVCSLLFGLRFSVEDKEFOAVRAAGTLLGVSSQGGQTYEMPSWFLRPLPGHKOLLH 239  
 Db 181 NVVCSLVFGLRPLVDDKEFOAVIAQASGTLGSSPWQAYEMPSWLLQPLPGHPTQLQH 240  
 QY 240 HVSTLAFTVRQVQHQGLDASGPARDLVDAFLKMAQEEQNGTFTNKNMLMTVYL 299  
 Db 241 HLGTLAFTVQVQKHGRFTSGPARDVDAFLKMAQEKQDPGTFTNKNMLMTVYL 300  
 QY 300 LFGATMTVSTVGYTLTLLMKYPHVQKWRVREELNRELGAQAPSLGDRTRLPYTDVLHE 359  
 Db 301 LFGATMTIGATIRYALLLLRYQVQVQVREELIQELGPGRAPSLSDRVLPTDVLHE 360  
 QY 360 AORLLALVPMGIPRTLMRTTRFRGYTLPGQTEVPFLLGSLIHDNPINKHPEFNPRFLD 419  
 Db 361 AORLLALVPMGMHTITRTTCFRGYTLPGKTEVPFLLGSLIHDNPINKHPEFNPRFLD 420  
 QY 420 ADRGRKHEAFLPFLSKRYCLGEGAKAELEFFFTILQAFSLSCPPDPTLSLKPVTUS 479  
 Db 421 EDORLRKHEAFLPFLSKRYCLGEGAKAELEFFFTILQAFSLSCPPDPTLSLKPVTUS 480  
 QY 480 GLFNIPPAFQOLVRPT 495  
 Db 481 GLFNIPDFQLRWPT 496

RESULT 4  
 Q29532

ID Q29532 PRELIMINARY; PRT; 491 AA.  
 AC Q29532; Q29515;  
 DT 01-NOV-1996 (TRENBLrel. 01, Created)  
 DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)  
 DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)  
 DE Cytochrome P450 2B-BX (BC 1.14.14.1).  
 OS Oryctolagus cuniculus (Rabbit).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
 OX NCBI\_TaxID=9986;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93349061; PubMed=8346920;  
 RA Ryan R., Grimm S.W., Kedzie K.M., Halpert J.R., Philpot R.M.;  
 RT "Cloning, sequencing, and functional studies of phenobarbital-  
 RT inducible forms of cytochrome P450 2B and 4B expressed in rabbit  
 RT kidney.";  
 RL Arch. Biochem. Biophys. 304:454-463 (1993).  
 RN [2]  
 RP SEQUENCE OF 2-491 FROM N.A.  
 RC STRAIN=NEW ZEALAND WHITE; TISSUE=Kidney;  
 RA Ryan R., Grimm S.W., Kedzie K.M., Halpert J.R., Philpot R.M.;  
 RL Submitted (OCT-1993) to the EMBL/GenBank/DBJ databases.  
 CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.  
 DR EMBL; S64259; AAB27705.1; -;  
 DR EMBL; L10912; AAA31432.1; -;  
 DR PIR; S35666; S35666.  
 DR HSSP; P00179; 1DT6.  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR GO; GO:0004497; F:monooxygenase activity; IEA.  
 DR GO; GO:0016712; F:oxidoreductase activity, acting on paired d. .; IEA.  
 DR GO; GO:0006118; P:electron transport; IEA.  
 DR InterPro; IPR001128; Cytochrome\_P450.  
 DR InterPro; IPR008068; EP450\_CYP2B.  
 DR Pfam; PF00067; P450; 1.  
 DR PRINTS; PR01685; EP450ICYP2B.  
 DR PRINTS; PR00385; P450.  
 DR PROSITE; PS00086; CYTOCHROME\_P450; 1.  
 KW Electron transport; Heme; Membrane; Monooxygenase; Oxidoreductase.  
 FT METAL 436 436 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).  
 SQ SEQUENCE 491 AA; 55732 MW; E3929C649FC66F17 CRC64;

Query Match 50.2%; Score 1313; DB 6; Length 491;  
 Best Local Similarity 50.7%; Pred. No. 4.7e-96;  
 Matches 24; Conservative 99; Mismatches 137; Indels 4; Gaps 3;

QY 7 WALLALALL--LILLTLALSGTRAGHLPPGPTPLPLGNLLQRLGALYGLMRLSKY 64  
 Db 3 FSLLLALLAFLAGLLLLFRGHFKAGRLPPGPTPLPLGNLLQMDRKLRSFRLREKY 62  
 QY 65 GPVFTYILGPWPVPVVLVQEAAREALGQAEFEFGRTVAMLEGTFDGHHGVFFSNGERW 124  
 Db 63 GDVFTVYLGSL-RPWVVLGTDALREALVDQAEAFSGRGKIAVVDPIFGYGVIFANGERN 121  
 QY 125 RQLRKFTMLALDLGKRGEGELIQAEARCLVETFGTGRFPDPSLLAQAATSNVVCS 184  
 Db 122 RALRRFSLATMRDPFGMKRSEVERIQEAEARCLVEELRKSFGALLDNTLLFHSITSNICS 181  
 QY 185 LLEGLRFSYEDKEFOAVRAAGTLLGVSSQGGQTYEMPSWFLRPLPGHKOLLHVVSL 244  
 Db 182 IVFGKFDYKDPVFLRLDLDFQSFSLISFSQVFLSGFLKHPFGTHQIRYRLQEI 241  
 QY 245 AAFVTRVQVQHQGLDASGPARDLVDAFLKMAQEEQNGTFTNKNMLMTVYLFPAGT 304  
 Db 242 NTFITQVSKHRATLDPNSP-RDFIDVYLLRMEKDSPPSEFHQNLILTVLSLFPAGT 300  
 QY 305 MTVSTTVGYTLTLLMKYPHVQKWRVREELNRELGAQAPSLGDRTRLPYTDVLHEAQRLL 364  
 Db 301 ETTSTTLRYGFLMLKYPHVTERVQKEIEQVIGSHRPPALDDRAKMPYTDVHEIQRLG 360  
 QY 365 ALVPMGIPRTLMRTTRFRGYTLPGQTEVPFLLGSLIHDNPINKHPEFNPRFLDADGRF 424  
 Db 361 DLIPFGVPHIVTDQFRGVIPIKNTVEFVLSSALHDPHYFKTPTNTPFNGHFLDANGAL 420

QY	425	RKHAFLPFLSKRVCLGEGIAKAEPLFFFTTILQAFSLSPCPDPTTSLKPTVSGLFNI	484
DB	421	KRNEGFMFSLKRVCLGEGIAKAEPLFFFTTILQAFSLSPCPDPTTSLKPTVSGLVNI	480
QY	485	PPAFOLQ	491
DB	481	PPSYQIR	487
RESULT 5			
Q8SQ67	Q8SQ67	PRELIMINARY;	PRT; 493 AA.
AC	Q8SQ67;	DT	01-JUN-2002 (TremBLrel. 21, Created)
DT	01-JUN-2002 (TremBLrel. 21, Last sequence update)		
DT	01-JUN-2003 (TremBLrel. 24, Last annotation update)		
DE	Cytochrome P450 2B22.		
GN	CYP2B22.		
OS	Sus scrofa (Pig).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.		
OX	NCBI_TaxID=9823;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Liver;		
RA	Kojima M.;		
RT	"Cloning of pig cytochrome P450 cDNAs and their expressions in		
RL	tissues";		
RL	Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.		
CC	-!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.		
DR	EMBL; AB052256; BAB85662.1; -.		
DR	GO; GO:0004497; F:monooxygenase activity; IEA.		
DR	GO; GO:0016712; F:oxidoreductase activity, acting on paired d. .; IEA.		
DR	GO; GO:0006118; P:electron transport; IEA.		
DR	InterPro; IPR001128; Cytochrome P450.		
DR	InterPro; IPR008088; EP450_CYP2B.		
DR	Pfam; PF00067; P450; 1.		
DR	PRINTS; PR01685; EP450ICYP2B.		
DR	PRINTS; PR00385; P450.		
DR	PROSITE; PS00086; CYTOCHROME P450; 1.		
KW	Heme; Monooxygenase; Oxidoreductase.		
SQ	SEQUENCE 493 AA; 56154 MW; 7f6322c41e4077e1 CRC64;		
Query Match 49.8%; Score 1301; DB 6; Length 493;			
Best Local Similarity 49.2%; Pred. No. 4.2e-95;			
Matches 240; Conservative 104; Mismatches 142; Indels 2; Gaps 2;			
QY	7	WALLLALALLLTALSGTRARGHLPPTPLPLGNLLQLRPGALYGLMELSKYGP	66
DB	4	FSAILLLLVGLVFLWGHPSQGHLPPTPLPLGNLLQLRPGALYGLMELSKYGP	63
QY	67	VFTYILGPNRPVVVLVGOAVREALGQAEFGSGRTVAMLEGTDFGCHGVFFSGNRWRQ	126
DB	64	VFTYILGP-RPVVMICGTEAIRREALVDQAEAFSGRGKIAVVEIFQYGVIFSGNRWKT	122
QY	127	LRKFTMLALRDLMGKREGBELIOAEARCLVETFTQGTGPPDPSLLAQAATNVVCSLL	186
DB	123	LRFSLATMRDGLGKRSVEERIQEAEQCLVBEELRSQGLQDPTTYFHSSTANICISV	182
QY	187	FLGRFSYEDKEFOAVRAAGGTLGVSSQGGQYEMFWSFLRPLPGPHKOLLHHVSLAA	246
DB	183	FGKRFAYRDEPFLQLDLPLFTLLTSSFPSSQLFELYSAFKVPFGSHRQIYKNLQOVNA	242
QY	247	FTVRVQOQHGNLDASGPARDLVDAFLKKMAQEONPGTEFTNKNMLMTVIYLLFAGTMT	306
DB	243	FGRSVEKHRETLDPSDP-RDLDSYLLRMDKESNSEFHEQNLLITVLSLFFACTET	301
QY	307	VSTTVGYTLMLLMKYPHQKVRBELNRELGAGQAPSLGDRTRLPYTDVLAHQAORLLAL	366
DB	302	TSTTLRYGFLMLKYPHITERIHKEIDRVIGSHRPPALDDRAKMPYMDAVIHEMQREGDL	361
QY	367	VPMGIPRTLMRTTFRGYTLPGQTEVPFLGSLILHDPNIFKHPEEPNDRFLDADGRFRK	426

Db	362	IPMGVPHVTXDTHERGVLLPKGIEVYILTSALHDPYFENPFAFNPDHFLDATGALKK	421
QY	427	HEAFLPFLSKRVCLGEGIAKAEPLFFFTTILQAFSLSPCPDPTTSLKPTVSGLVNI	486
DB	422	NDAFMPFSIGKRICLGEGLARTEFLFFFTTILQAFSLSPCPDPTTSLKPTVSGLVNI	481
QY	487	AFOLQVRP	494
DB	482	PYQIQFLP	489
RESULT 6			
Q8K027	Q8K027	PRELIMINARY;	PRT; 323 AA.
AC	Q8K027;	DT	01-OCT-2002 (TremBLrel. 22, Created)
DT	01-OCT-2002 (TremBLrel. 22, Last sequence update)		
DT	01-JUN-2003 (TremBLrel. 24, Last annotation update)		
DE	Similar to RIKEN CDNA 1200011C15 gene (Fragment).		
GN	CYP2S1 OR 1200011C15RIK.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Colon;		
RA	Strausberg R.;		
RL	Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.		
CC	-!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.		
DR	EMBL; BC034202; AAH34202.1; -.		
DR	MGD; MGI:1321384; Cyp2s1.		
DR	GO; GO:0004497; F:monooxygenase activity; IEA.		
DR	GO; GO:0006118; P:electron transport; IEA.		
DR	InterPro; IPR001128; Cytochrome P450.		
DR	Pfam; PF00067; P450; 1.		
DR	PRINTS; PR00385; P450.		
KW	Heme; Monooxygenase; Oxidoreductase.		
NON TER	1		
SQ	SEQUENCE 323 AA; 36065 MW; 71BD0875E35C0E48 CRC64;		
Query Match 49.7%; Score 1300; DB 11; Length 323;			
Best Local Similarity 76.1%; Pred. No. 2.8e-95;			
Matches 242; Conservative 37; Mismatches 39; Indels 0; Gaps 0;			
QY	178	TSNVCSLLFLGRFSYEDKEFOAVRAAGGTLGVSSQGGQYEMFWSFLRPLPGPHKQL	237
DB	1	TSNVCSLVFGIRLIPYDDKEFOAVIQASGTLIGISSPMQAYEMFWSLLQPLPGPHQTOL	60
QY	238	LHHVSTLAFTVRVQOQHGNLDASGPARDLVDAFLKKMAQEONPGTEFTNKNMLMTVI	297
DB	61	QHHLGLTAFTTIOQVQKHQGFQTSGPARDVDVDAFLKKMAQEONPGTEFTNKNMLMTVT	120
QY	298	YLLFAGTMTVSTTVGYTLMLLMKYPHQKVRBELNRELGAGQAPSLGDRTRLPYTDVLA	357
DB	121	YLLFAGTMTIGATIRYALLLLLRYPOVQQRVREELIQELGPGRAPSLSDRVRLPYTDVLA	180
QY	358	HEAQRLLALVPMGIPRTLMRTTFRGYTLPGQTEVPFLGSLILHDPNIFKHPEEPNDRF	417
DB	181	HEAQRLLALVPMGMPHTITRTTCFRGYTLPGTEVPFLGSLILHDPVAFVQNGFPHGRF	240
QY	418	LDADGRFRHAEPLPSLGRVCLGEGIAKAEPLFFFTTILQAFSLSPCPDPTTSLKPT	477
DB	241	LDEGRRLRHAEPLPSLGRVCLGEGIAKAEPLFFFTTILQAFSLSPCPDPTTSLKPT	300
QY	478	VSLFNIPPAFOLQVRPT	495
DB	301	ISGLFNIPDPQLRVWPT	318
RESULT 7			
Q29516			



ID	Q29516	PRELIMINARY;	PRT;	491 AA.
AC	Q29516;			
DT	01-NOV-1996	(TREMBLrel. 01, Created)		
DT	01-NOV-1996	(TREMBLrel. 01, Last sequence update)		
DT	01-JUN-2003	(TREMBLrel. 24, Last annotation update)		
DE	Cytochrome P-450 (P450IIB)			
OS	Oryctolagus cuniculus (Rabbit).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.			
NCBI	TaxID=9986;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
EX	MEDLINE=88163620; PubMed=2831964;			
RA	Komori M., Imai Y., Tsunasawa S., Sato R.;			
RT	"Microheterogeneity in the major phenobarbital-inducible forms of			
RT	rabbit liver microsomal cytochrome P-450 as revealed by nucleotide			
RT	sequencing of cloned cDNAs";			
RL	Biochemistry 27:73-80(1988).			
CC	-I- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.			
DR	EMBL; M18820; AAA31434.1; -.			
DR	PIR; A27717; A27717.			
DR	HSSP; P00179; 1DT6.			
DR	GO; GO:0004497; P:monooxygenase activity; IEA.			
DR	GO; GO:0016712; E:oxidoreductase activity, acting on paired d. . .; IEA.			
DR	GO; GO:0006118; E:electron transport; IEA.			
DR	InterPro; IPR001128; Cytochrome P450.			
DR	InterPro; IPR008068; EP450_CYP2B.			
DR	Pfam; PF00067; P450; 1.			
DR	PRINTS; PR01685; EP450ICYP2B.			
DR	PRINTS; PR00385; P450.			
DR	PROSITE; PS00086; CYTOCHROME P450; 1.			
KW	Heme; Monooxygenase; Oxidoreductase.			
SQ	SEQUENCE 491 AA; 55686 MW; 14B5F2B1BF783A8F CRC64;			
	Query Match 49.6%; Score 1298; DB 6; Length 491;			
	Best Local Similarity 51.0%; Pred. No. 7.3e-95;			
	Matches 249; Conservative 98; Mismatches 135; Indels 6; Gaps 5			
Qy	7	WALLALLALL--LLLTALSGTRAGHLPPGPTPLPLGLNLLQL--RPGALYGLMLRLSKK 63		
Db	3	FSLLLLLAFLAGLLLLLLLRGHPKAGRLPPGPPPLPVGLNLLQMDRKGLLS--FTQLREK 61		
Qy	64	YGPVFTIYLGPRPVRVVLVGOEAVREALGGOAEFSGRGTVAMLEGTGDGHGVFSSNGER 123		
Db	62	YGDVFTVVLGS--RPVVVL--CGTDAIREALVDQAEAFSGRGKTAVLDPITFGVGVTFANGER 120		
Qy	124	WRQLRKETMALRDGLMGKRGSEBELIQAEACLVETFOGTGPRPDPDLLLAQAATSNVVC 183		
Db	121	WRALRRESLATMRDFGMGKRSVEERIOPEARCLVEELRKSGALLDNTLLFHSVTSNII 180		
Qy	184	SLAFLGRFSYEDKEFQAVVRAAGTLLGVSSOGQTYEMFSWFLRPLGPHKQLLHVST 243		
Db	181	SIVFGKRFYKDPVFLRLDLFFQSFSITSFSQVFLFSGFLKHFPGTHRQIYRNQ 240		
Qy	244	LAALTFRVQOQHGNLDASGPARDLVDAFLIKMAEQEONPGTEFTKNMLMTVIYLLFAG 303		
Db	241	INTFIGQSVHERHETLDPNSP--RDFIDVYLLRMKDKSDPSSEFHHQMLLIUTVTLFPAG 299		
Qy	304	TMTVSTTVGYTLLLLMKYPHVQKWRBELNRELGAQAPSLGDRTRLPYTDVAVLHEAQRL 363		
Db	300	TETTSTTYRGFLLLKYPHVTREVQKEIQEVIGSHRPPALDDRAKMPYTDVAVIHEIQL 359		
Qy	364	LALVPMGIPRTLMTTFTRFGYTLPCQGTVEFPLLSILHDPNI FKHPBEFNPRFLDADGR 423		
Db	360	GDILPLGVPMHVTKOTDFRGVPIPKNTEVFVLSSALHDPFRYFETPNTFNGHFLDADGA 419		
Qy	424	FRXHEAFPLPSLGRKVCGLGKAAELFLPFTTILQAFSLSESPCPDTLSLKPTVSGLFN 483		
Db	420	LXKIEGMPFSLGKRICLGEIARTLELFPFTTILQNFSTASPPVEDIDLTTPRESGVGN 479		
Qy	484	IPPAFQLQ 491		
Db	480	VPPSYOIR 487		

```

RESULT 8
Q9WUD0
ID Q9WUD0 PRELIMINARY; PRT; 491 AA.
AC Q9WUD0;
DT 01-NOV-1999 (T-EMBLrel. 12, Created)
DT 01-NOV-1999 (T-EMBLrel. 12, Last sequence update)
DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
DE Cytochrome P450 2B10 related protein.
GN CYP2B20.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RP STRAIN=BALB/c; TISSUE=Liver;
RC MEDLINE=99240492; PubMed=10222226;
RX Marc N., Damon M., Fautrel A., Guillozo A., Corcos L.;
RT "Isolation of a cyp2b10-like cDNA and of a clone derived from a
RT cyp2b10-like pseudogene."
RL Biochem. Biophys. Res. Commun. 258:11-16(1999) .
RN [2]
SEQUENCE FROM N.A.
RP STRAIN=C57BL/6J; TISSUE=Pancreas;
RC MEDLINE=22354683; PubMed=12466851;
RX The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
CC Nature 420:563-573(2002) .
CC 1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR ENBL; AF128849; RAD28466.1; -.
DR ENBL; AK028103; BAC25748.1; -.
DR HSSP; P00179; 1DT6.
DR MGD; MGI:1202389; Cyp2b20.
DR GO; GO:0004497; F-monooxygenase activity; IEA.
DR GO; GO:0016712; F-oxidoreductase activity, acting on paired d. . .; IEA.
DR GO; GO:0006118; P- electron transport; IEA.
DR InterPro; IPR001128; Cytochrome P450.
DR InterPro; IPR008068; Cytochrome P450.
DR Pfam; PF00067; p450; 1.
DR PRINTS; PR01685; EP450ICVP2B.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME P450; 1.
KW Heme; Monooxygenase; Oxidoreductase.
SQ SEQUENCE 491 AA; 55870 MW; A1338E108D77025E CRC64;

```

Query Match	49.6%	Score 1296;	DB 11;	Length 491;
Best Local Similarity	50.8%;	Pred. No. 1.le-94;		
Matches	245;	Conservative 96;	Mismatches 139;	Indels 2; Gaps 2;
QY	9	LLALALLLLLTALSGTRARGHLPGPPTPLPLGNLLQLRPGALYSGLMRLSKKYPGVF	68	
Db		:       :       :       :       :       :       :       :       :       :		
	7	LLTALVGFLLLLARGHPKSRGNFPDPGRPLPLGNLLQMDRGCLLKSFIQLREKYGDVF	66	
QY	69	TIYLGPMRPVVLVGOAEVREALGGQAEFSRGRTVAMLEGTDGHHGVFSNGERWRQLR	128	
Db		:       :       :       :       :       :       :       :       :       :		
	67	TVHGLP-RDVMVLCGGDTTREALVGQAEAFSGRGTVAVVEPTFEKEYGVI-FANGERWKTLR	125	
QY	129	KFTMLALRDLGKMGKEGELIOAEARCLVETFGTREGRPDPSPSLLAQAATSNVVCSLIFG	198	
Db		:       :       :       :       :       :       :       :       :       :		
	126	RSLATMRDFFGMGKSVERRIQEACLVVEELRKSQGAFLDPTFLFCQTANICISIVFG	185	
QY	189	LFSEYDEKFEQAVVRAAGTLLGVSSGGQTYEMFSWFLRPLPGPHKQLLHHVSTLLAAFT	248	
Db		:       :       :       :       :       :       :       :       :       :		
	186	ERFEYTDQFLRLLELYFTLSLISFSQMFELFSGFLKYPFGAHRQISKNLQELLDIYI	245	
QY	249	VZQVQOQHGNLDASGPARDLVDAFLIKMAQEQNPCTETFTNKNMLMTVIYLLFAGTMTVS	308	
Db		:       :       :       :       :       :       :       :       :       :		
	246	GHSVEKRALDPSVP-RDFIDIYLLRMEKESKNOHTPEFHONLMMSVLSLFPAGTETSS	304	

Mon Sep 20 09:09:28 2004

```
Qy 309 TTGTYTLLLMKYPHVQKRWREELNRELGAGQAPSLGDRTRLPYTDVHLHEAQRLLALVP 368
Db 305 TTRYGFLMLKYPHVAEKVQKEIDQVIGSHRLPTLDDRTKMPYTDVAIHEIQRFSDLP 364
Qy 369 MGIPRLMTRTFRGYTLPQGTVEVPLGLSLHLDENIFKHPEFNPDRLDADGRFKHE 428
Db 365 IGVPRVTKMTFRGYLPLPKTEVYPIUSSALHDPQYEQPDNPDHLDANGALKKSE 424
Qy 429 AFLPSELGKVCGLGSLAKAEFLFFTTILQAFSLSPCPPDLSLKFTVSLGNIPPAF 488
Db 425 AFLPESTGKICLGSIAARNELFFTSILQNFVASHVAPKIDILTPEKSGIGKIPPT 484
Qy 489 QL 490
Db 485 QI 486

RESULT 9
ID Q9Q250 PRELIMINARY; PRT; 494 AA.
AC Q9Q250;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Cytochrome P-450.
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Mesocricetus.
ON NCBI_TaxID=10036;
RX TISSUE=Liver;
RC SEQUENCE FROM N.A.
RA Tohkin M., Kurose K., Isozaki E., Fukuhara M.;
RT "Molecular cloning, heterologous expression, and characterization of a
RL novel member of CYP2A in the Syrian hamster."
RL Biochim. Biophys. Acta 1446:438-442 (1999).
CC -I- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR EMBL; D86952; BAA85463.1; -.
DR HSSP; P00179; 1DT6.
DR GO; GO:0004497; F:monooxygenase activity; IEA.
DR GO; GO:0016712; F:oxidoreductase activity, acting on paired d. . .; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR001128; Cytochrome P450.
DR InterPro; IPR008067; EP450_CYP2A.
DR Pfam; PF00067; P450; 1.
DR PRINTS; PR01684; EP450ICYP2A.
DR PROSITE; PR00385; P450.
DR Heme; Monooxygenase; Oxidoreductase.
KW SEQUENCE 494 AA; 56544 MW; A0A37193E2BC5849 CRC64;

Query Match 48.2%; Score 1261.5; DB 11; Length 494;
Best Local Similarity 50.0%; Pred. No. 5.9e-92;
Matches 247; Conservative 92; Mismatches 146; Indels 9; Gaps 4;

Qy 1 MEATGTWALLLALLLLTALSST----RARGHLPPGPTPLPLGNLLQLRPGALYS 56
Db 1 MLASG---LLLVTVLAFVLVILMSVWKQKLSGKLPPGPTPLPFIGNYLQNLTEQMYNS 57
Qy 57 LMLSKYGPVFTTYLGFWRPVVVLVQBAVREALGQAEFSGRGTVAMLEGTFFDGHV 116
Db 58 LMKISRYGPVFTTHLGP-RPIVVLGQEAVKALVDQAEFSGRGQATFDMLFKQYGV 116
Qy 117 FNSNGERWROLRKFTMLALRDLGMGKREGBELIAQEARCLVETPQGTGRRPDPSSLIAQ 176
Db 117 AFSSGERAKQRRFSIATLDRFVGKGIERIEEAGFLLEAFKXNGALIDTFFYLSR 176
Qy 177 ATSNVCSLLFGLRFSYEDKEFOAVVRAAGTLLGVSSQGGQTVEMFSWFLRPLPGPHKQ 236
Db 177 TVSNVISSIVFGDRFDEKEDKFLSLLRMMLGSFQGTGTSTGQLYEMFSSVMKHLPGPQQQ 236
```

```
Qy 237 LLHHVSTLAAFTVROVQOQGNLDASGPARDLVDALFLKMAQEEQNPGTEFTNKMLMTV 296
Db 237 AFKEIQLGLEDITTKYEQNQRTLDNSP-RDFIDSFLIRMKEEKKNPTEFFMKNLVLT 295
Qy 297 IYLLFAGTMTVSTTVGYTLLLLMKYPHVQKRWREELNRELGAGQAPSLGDRTRLPYTDV 356
Db 296 LNLFFAGTETVSTTVRYGYFLLLMKHPDVEAKVQEEIDRVIGKNRQPKYEDRLKMPYTEAV 355
Qy 357 LHEAQRLLALVPMGIPRLMTRTFRGYTLPQGTVEVPLGLSLHLDENIFKHPEFNPDRL 416
Db 356 IHEIQRFSDLPMTGLMGLARRVTKDTKFRGFFIPKGTVEVPMGLSVLRDPPKFFSNPKDNFPOH 415
Qy 417 FLDDAGFRKHEAFLPFSILGKRVCLGEGIAKAEFLFFTTILQAFSLSPCPPDLSLKP 476
Db 416 FLDDKGQFKKNDAFVPPSIGKEYCFEGELARMEFLFLNLLQNFHLRSFQAPQDIDVSP 475
Qy 477 TVSGLEFNIPPAFQL 490
Db 476 RLVGFEATIPPNTYM 489

RESULT 10
ID Q64584 PRELIMINARY; PRT; 491 AA.
AC Q64584;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Rat cytochrome P-450b (phenobarbital-inducible).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
ON NCBI_TaxID=10116;
RX SEQUENCE FROM N.A.
RA MEDLINE=85234490; PubMed=2989270;
RA Suwa Y., Mizukami Y., Sogawa K., Fujii-Kuriyama Y.;
RT "Gene structure of a major form of phenobarbital-inducible cytochrome
RL P-450 in rat liver."
RL J. Biol. Chem. 260:7980-7984 (1985).
CC -I- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR EMBL; L00320; AAA41046.1; -.
DR EMBL; M11251; AAA41046.1; JOINED.
DR EMBL; L00313; AAA41046.1; JOINED.
DR EMBL; L00314; AAA41046.1; JOINED.
DR EMBL; L00315; AAA41046.1; JOINED.
DR EMBL; L00316; AAA41046.1; JOINED.
DR EMBL; L00317; AAA41046.1; JOINED.
DR EMBL; L00318; AAA41046.1; JOINED.
DR EMBL; L00319; AAA41046.1; JOINED.
DR PIR; A00176; OART2B.
DR HSSP; P00179; 1DT6.
DR GO; GO:0004497; F:monooxygenase activity; IEA.
DR GO; GO:0016712; F:oxidoreductase activity, acting on paired d. . .; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR001128; Cytochrome P450.
DR InterPro; IPR008068; EP450_CYP2B.
DR Pfam; PF00067; P450; 1.
DR PRINTS; PR01685; EP450ICYP2B.
DR PROSITE; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME P450; 1.
DR Heme; Monooxygenase; Oxidoreductase.
KW SEQUENCE 491 AA; 56047 MW; E9144D72EF522F0C CRC64;

Query Match 48.1%; Score 1257; DB 11; Length 491;
Best Local Similarity 48.8%; Pred. No. 1.3e-91;
Matches 235; Conservative 100; Mismatches 145; Indels 2; Gaps 2;

Qy 9 LLALALLLLTALSSTTRARGHLPPGPTPLPLGNLLQLRPGALYSGLMRLSKYGPVF 68
Db 7 LLALLVGFLLLVGRGPKSGNFPFGPRLPLGNLLQLDRGGLLNSFMQLREKYGVDF 66
Qy 69 TIVLGPWRPVVVLVQBAVREALGQAEFSGRGTVAMLEGTFFDGHVFSNGERWRLR 128
```

```

Db 67 TVHLGP-RPVNMLCGTDTIKALVGPEDFSRGITAVIEPIKEXGVIFANGERKALR 125
QY 129 KFTMLALRDLGMKREGEELIOAEARCLVETTCGTGRFPDPSLLLAQAATSNVVCSLF 188
Db 126 RFLATMRDGMGKRSVEERIQEEAQCLVEELRKSQAGPLDPTFLQOCITANICISIVFG 185
QY 189 LRFSEYDEKFOAVRAAGCTLLGVSSQGGOTYEMESWFLRPLPGPHKQLLHVSTLAFT 248
Db 186 BRFDYTDQFLRLLELFFRFSLLSSFSQVEFFSGFLKYPGAGHRSKRLQELDYI 245
QY 249 VRQVOHQGNLDASGARDLVDAFLKMAQEONFGTEFTNKNMLMTVIYLLFAGTMTVS 308
Db 246 GHIVEKRAATLPSAP-RDFIDTYLLRMEKEKSNHTEFHENLMLISLLSFFAGTETS 304
QY 309 TTGVGTYLLMLKYPHVQKWRVRELNRELGAGQAPSLGDRTRLPYTDVLAHEAQRLLALVP 368
Db 305 TLLRYGFLMLKYPHVAEKVQREIDHVGISDRPPSLDDRRPKMPYTEAVIYEIQFSDLP 364
QY 369 MGIPRTLMRTFRGYTLPGQTEVPPLGSLIHDNFIKHPPEFNPDRFLDADGRFKHE 428
Db 365 IGPHRVTKDMWERYLLPKTEVYPIRSSALHDPOYFDHDSFNPHEFLDVGALKKSE 424
QY 429 AFLPFSIGKRVCLGEGLAKEAELFFFTTILQAFSLESPPDPTLSLKPTVSGLENIPPAF 488
Db 425 AFMPFSTGKHICLGEIARNELEFFFTTILQNFSSVSHLAPKDIIDLTPEKSGIKIPPT 484
QY 489 QL 490
Db 485 QI 486

```

```

RESULT 11
Q80XG6 PRELIMINARY; PRT; 491 AA.
ID AC Q80XG6
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Cytochrome P450, family 2 subfamily B polypeptide 29.
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Mesocricetus.
OX NCBI_TaxID=10036;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Harderian gland;
RA Alvarez J., Vinas J., Dominguez P.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY250712; AAP04405.1; -.
DR GO; GO:0016712; P:oxidoreductase activity, acting on paired d. .; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR001128; Cytochrome P450.
DR InterPro; IPR008068; EP450_CYP2B.
DR Pfam; PF00067; p450; 1.
DR PRINTS; PR01685; EP450ICYP2B.
DR PROSITE; PS00385; P450.
DR PROSITE; PS00086; CYTOCHROME_P450; 1.
SQ SEQUENCE 491 AA; 56298 MW; 34AD1B044DC45850 CRC64;

```

```

Query Match 48.0%; Score 1255; DB 11; Length 491;
Best Local Similarity 49.8%; Pred. No. 1.9e-91;
Matches 240; Conservative 91; Mismatches 149; Indels 2; Gaps 2;
QY 9 LLLALLLLLTALSTRAGHLPQGTPLPLGNLLQLRPGALYSGMLSKKYGVPV 68
Db 7 LLLTLLSFLVLLVGVGPKTRGHLPGPRPLPLGNLLQLDRGLNSFRPFKEKGDVP 66
QY 69 TIVLGPWRPVVLVQAEVREALGGQAEBSRGVTVAMLEGTDGPHGVFSNGERWRQR 128
Db 67 TLHLSG-KPVVMVYGEATREALVDQAEAFSGRTTIAVLKPTMDYGVIAANGRWKTLR 125

```

```

QY 129 KFTMLALRDLGMKREGEELIOAEARCLVETTCGTGRFPDPSLLLAQAATSNVVCSLF 188
Db 126 RFLATMRDGMGKRSVEERIQEEAQCLVEELRKSQAGPLDPTFLQOCITANICISIVFG 185
QY 189 LRFSEYDEKFOAVRAAGCTLLGVSSQGGOTYEMESWFLRPLPGPHKQLLHVSTLAFT 248
Db 186 BRFDYTDQFLRLLELFFRFSLLSSFSQVEFFSGFLKYPGAGHRSKRLQELDYI 245
QY 249 VRQVOHQGNLDASGARDLVDAFLKMAQEONFGTEFTNKNMLMTVIYLLFAGTMTVS 308
Db 246 GHIVEKRAATLPSAP-RDFIDTYLLRMEKEKSNHTEFHENLMLISLLSFFAGTETS 304
QY 309 TTGVGTYLLMLKYPHVQKWRVRELNRELGAGQAPSLGDRTRLPYTDVLAHEAQRLLALVP 368
Db 305 TLLRYGFLMLKYPHVAEKVQREIDHVGISDRPPSLDDRRPKMPYTEAVIYEIQFSDLP 364
QY 369 MGIPRTLMRTFRGYTLPGQTEVPPLGSLIHDNFIKHPPEFNPDRFLDADGRFKHE 428
Db 365 IGPHRVTKDMWERYLLPKTEVYPIRSSALHDPOYFDHDSFNPHEFLDVGALKKSE 424
QY 429 AFLPFSIGKRVCLGEGLAKEAELFFFTTILQAFSLESPPDPTLSLKPTVSGLENIPPAF 488
Db 425 AFMPFSTGKHICLGEIARNELEFFFTTILQNFSSVSHLAPKDIIDLTPEKSGIKIPPT 484
QY 489 QL 490
Db 485 QI 486

```

```

RESULT 12
Q64463 PRELIMINARY; PRT; 491 AA.
ID AC Q64463
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Testosterone 16a-hydroxylase type a.
GN CYP2B9 OR 16AOH-A.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91146586; PubMed=1997326;
RA Lakso M., Masaki R., Noshiro M., Negishi M.;
RT "Structures and characterization of sex-specific mouse cytochrome P-
RT 450 genes as members within a large family. Duplication boundary and
RT evolution."
RL Eur. J. Biochem. 195;477-486(1991).
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR EMBL; M60273; AAA03648.1; -.
DR EMBL; M60267; AAA03648.1; JOINED.
DR EMBL; M60268; AAA03648.1; JOINED.
DR EMBL; M60269; AAA03648.1; JOINED.
DR EMBL; M60270; AAA03648.1; JOINED.
DR EMBL; M60271; AAA03648.1; JOINED.
DR EMBL; M60272; AAA03648.1; JOINED.
DR PIR; I84735; I84735.
DR HGSP; P00179; ID76.
DR MGD; MGI:88600; Cyp2b9.
DR GO; GO:0004497; F:monooxygenase activity; IEA.
DR GO; GO:0016712; F:oxidoreductase activity, acting on paired d. .; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR001128; Cytochrome P450.
DR InterPro; IPR008068; EP450_CYP2B.
DR Pfam; PF00067; p450; 1.
DR PRINTS; PR01685; EP450ICYP2B.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME_P450; 1.
KW Heme; Monooxygenase; Oxidoreductase.
SQ SEQUENCE 491 AA; 55740 MW; C1E790A7DD1A7298 CRC64;

```

Query Match 47.7%; Score 1247; DB 11; Length 491;  
 Best Local Similarity 48.8%; Pred. No. 8.4e-91;  
 Matches 235; Conservative 99; Mismatches 146; Indels 2; Gaps 2;

QY 9 LLLALLLLTLLSGLTRARGHLPGPTPLPLGGLNLLQRLPGALYSGMLRLSKYGVVF 68  
 DB 7 LLLAVLLLSLLVVRGHAKIHGLPPGPHPLPLGGLNLLQMDRGKLLKCFIQLOERKGDVF 66

QY 69 TIYLGPMRPVVVLVGOEAVREALGQAEFFSGRGTVAMLEGTFDGHGVFFSNGERWQLR 128  
 DB 67 TVHLGP-RPVVVLGTTQITREALVDHAEAFSGRTIAAAQLVMQDYGIFFASGQRWKTIR 125

QY 129 KFTMLARDLGMKREGEELIQAEARCLVETFGTEGRFPDPSLLLAQATSNVCSLLFG 188  
 DB 126 RFSLATMKEFGMGKRSVEERIKEEAQCVELKKGQVPLDPTFLFCQTANILCSIVFG 185

QY 169 LFFSYEDKEFOAVRAAGTLLGVSSOGQTYEMFSWFLRPLPGPHKQLLHHVSTIAAFT 248  
 DB 186 ERFDYTDQFHLNLMYKIFSLSSFSQMFELSGFLKYPGVHRQIVKKQBELLDVI 245

QY 249 VRQVOQHQNLDASGPARDLVDAFLKMAQEQNPGTEFTNQMLMTVIYLLFAGTMVTS 308  
 DB 246 AHSVEKHATLPSAP-RDYIDTYLLRMEKEKSNHNTFHHQNLMMSVLSLFFAGTETTS 304

QY 309 TTVGYTLLMLKYPHVQKVRRELNRELGAQPSLGDRTLPYTDVAVLHEAQRLLALVP 368  
 DB 305 ATLHYGVLLMLKYPHVTEKQVEIDQVIGSHRLPTLDDRTKMPYTDVAVLHEIQRFSDLPV 364

QY 369 MGIPRLMTTRPRGYTLPGTGVFFPLGSLIHDNPIFKHPEEFNPDRLDADGRFKHE 428  
 DB 365 IGLPHKVIKDTLFRGYLLPKNTEYFVLSALHDPQYFQPKNFENHEFLDANGALKKCE 424

QY 429 AFLPFLSGKRVCLGEGAKAELFFFTTILQAFSLSPCPDPLSLKPTVSGLFNPPAF 488  
 DB 425 AFLPFTGKRIKLGESIAELNRELFFFTTILQAFSLSPCPDPLSLKPTVSGLFNPPAF 484

QY 489 QL 490  
 DB 485 QI 486

RESULT 13  
 Q9WV19 PRELIMINARY; PRT; 494 AA.  
 AC Q9WV19  
 DT 01-NOV-1999 (T-EMBLrel. 12, Created)  
 DT 01-NOV-1999 (T-EMBLrel. 12, Last sequence update)  
 DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)  
 DE Olfactory-specific steroid hydroxylase.  
 GN CYP2G1.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Olfactory neuroepithelium;  
 RX MEDLINE=97288345; PubMed=9143323;  
 RA Hua Z., Zhang Q.Y., Su T., Lipinskas T.W., Ding X.;  
 RT "cDNA cloning, heterologous expression, and characterization of mouse  
 CYP2G1, an olfactory-specific steroid hydroxylase.";  
 Arch. Biochem. Biophys. 340:208-214(1997).  
 RL -/- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.  
 CC EMBL; L81171; AAD45178.1; -;  
 DR HSP; P00179; 1DT6.  
 DR MGP; MG1109612; Cyp2g1.  
 DR GO; GO:0004497; F:monooxygenase activity; IEA.  
 DR GO; GO:0016712; F:oxidoreductase activity, acting on paired d. . .; IEA.  
 DR GO; GO:0006118; P:electron transport; IEA.  
 DR InterPro; IPR001128; Cytochrome P450.  
 DR Pfam; PF00067; P450; 1.  
 DR PRINTS; PR01684; EP450ICYP2A.

DR PRINTS; PR00385; P450.  
 DR PROSITE; PS00086; CYTOCHROME P450; 1.  
 KW Heme; Monooxygenase; Oxidoreductase.  
 SQ SEQUENCE 494 AA; 56819 MW; 63D83B09DB08AB9F CRC64;

Query Match 47.6%; Score 1243.5; DB 11; Length 494;  
 Best Local Similarity 48.3%; Pred. No. 1.6e-90;  
 Matches 235; Conservative 96; Mismatches 153; Indels 3; Gaps 3;

QY 5 GTWALLLAL-LLLTLLSGLTRARGHLPGPTPLPLGGLNLLQRLPGALYSGMLRLSKK 63  
 DB 5 GAFSIFMALCSCLLIILAWKRTSGKKLPPGPTPIPLGNFLQVTRTATFOSFQKLOK 64

QY 64 YGEVETIYLGPMRPVVVLVGOEAVREALGQAEFFSGRGTVAMLEGTFDGHGVFFSNGER 123  
 DB 65 YGSVTVVFGP-RPVVVLGCGHAEVKEALVDQADDFSGRGEMPTLEKNFGQYGLALSNGR 123

QY 124 WQLAKFTMLARDLGMKREGEELIQAEARCLVETFGTEGRFPDPSLLLAQATSNVVC 183  
 DB 124 WKILRRFSLTVLRNFGMGKRSIEERIOEEAGVLLLELHKVKGAPIDPTIYLSRTVSNVIC 183

QY 184 SLLFGLRFSYEDKEFOAVRAAGTLLGVSSOGQTYEMFSWFLRPLPGPHKQLLHHVST 243  
 DB 184 SVVFGKRFYDQQRQSLMRMINESFVMSKPAQLYDMYKVMQYFPGRHVLYNLIED 243

QY 244 LAAFTVRQVQOQHQNLDASGPARDLVDAFLKMAQEQNPGTEFTNQMLMTVIYLLFAG 303  
 DB 244 LKDFIASRVKINEASDPSPN-RDFIDCEFLMKHQKSDPHTFENLKNLVLTLNLF 302

QY 304 TMTVSTTVGYTLLMLKYPHVQKVRRELNRELGAQPSLGDRTLPYTDVAVLHEAQRLL 363  
 DB 303 TETVSSTLYRGYFLLKYPEVEAKIHEEINQVITGTHRTPRVDDRAKMPYTDVAVLHEIQRL 362

QY 364 LALVPMGIPRTLMRTTRFRGYTLPGTGVFFPLGSLIHDNPIFKHPEEFNPDRLDADGR 423  
 DB 363 TDIVPUGVPHNVTRDTHFRGYLLPKGTVDYPLFGSVLKDPKYFRYPDAFPQHFQDQGR 422

QY 424 FRKHEAFLPFLSGKRVCLGEGAKAELFFFTTILQAFSLSPCPDPLSLKPTVSGLFN 483  
 DB 423 FKNDADFVVFSSGKRICVGEALARMELFYFTSILQRFSLRSLVPPADIDIAHKISGFGN 482

QY 484 IPPAFOL 490  
 DB 483 IPPVVEL 489

RESULT 14  
 Q8SQ68 PRELIMINARY; PRT; 494 AA.  
 AC Q8SQ68  
 DT 01-JUN-2002 (T-EMBLrel. 21, Created)  
 DT 01-JUN-2002 (T-EMBLrel. 21, Last sequence update)  
 DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)  
 DE Cytochrome P450 2A19.  
 GN CYP2A19.  
 OS Sus scrofa (Pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
 OX NCBI\_TaxID=9823;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver;  
 RA Kojima M.;  
 RT "Cloning of pig cytochrome P450 cDNAs and their expressions in  
 tissues.";  
 Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.  
 CC -/- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.  
 DR EMBL; AB052255; BAB85661.1; -;  
 DR GO; GO:0004497; F:monooxygenase activity; IEA.  
 DR GO; GO:0016712; F:oxidoreductase activity, acting on paired d. . .; IEA.  
 DR GO; GO:0006118; P:electron transport; IEA.  
 DR InterPro; IPR001128; Cytochrome P450.  
 DR Pfam; PF00067; P450; 1.  
 DR PRINTS; PR008067; EP450\_CYP2A.



***This Page Blank (uspto)***



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2004, 09:14:40 ; Search time 25 Seconds  
(without alignments)  
1049.734 Million cell updates/sec

Title: US-10-669-693-2

Perfect score: 2615

Sequence: 1 MEATGTWALLALALLLLT.....PPAFQLQVRPTDLHSTTQTR 504

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1319.5	50.5	494	1	CPG1_RABIT
2	1300	49.7	491	1	CPB4_RABIT
3	1296	49.6	491	1	CPB2_RAT
4	1296	49.6	491	1	CPB5_RABIT
5	1290	49.3	491	1	CPB1_RAT
6	1280.5	49.0	500	1	CPBA_MOUSE
7	1262	48.3	494	1	CPBB_CANFA
8	1244	47.6	491	1	CPB9_MOUSE
9	1243.5	47.6	494	1	CPG1_RAT
10	1238	47.3	491	1	CPB6_HUMAN
11	1237	47.3	494	1	CPA5_MOUSE
12	1236.5	47.3	492	1	CPBJ_MOUSE
13	1235	47.2	494	1	CPA4_MOUSE
14	1234.5	47.2	494	1	CPA3_RAT
15	1214.5	46.4	494	1	CPAD_HUMAN
16	1212.5	46.4	494	1	CPA6_HUMAN
17	1210.5	46.3	492	1	CPBC_RAT
18	1203.5	46.0	494	1	CPAA_RABIT
19	1198.5	45.8	494	1	CPAB_RABIT
20	1199	45.5	494	1	CPBK_MOUSE
21	1176.5	45.0	494	1	CPA7_HUMAN
22	1172	44.8	491	1	C2F3_CAPHI
23	1168	44.7	491	1	C2F2_MOUSE
24	1163.5	44.5	494	1	CPCN_RAT
25	1160	44.4	491	1	CPB3_RAT
26	1158	44.3	493	1	CPA9_MBSAU
27	1150.5	44.0	492	1	CPAC_MOUSE
28	1145	43.8	491	1	C2F4_RAT
29	1138.5	43.5	492	1	CPAL_RAT
30	1108	42.4	491	1	C2F1_HUMAN
31	1107	42.3	491	1	CPH2_CHICK
32	1099.5	42.0	494	1	CPA8_MBSAU
33	1098	42.0	487	1	CPCG_RABIT

34 1095.5 41.9 490 1 CPZ4\_MOUSE  
35 1093 41.8 487 1 CPC5\_RABIT  
36 1092.5 41.8 490 1 CPC7\_RAT  
37 1091.5 41.7 492 1 CPA2\_RAT  
38 1085.5 41.5 490 1 CPC9\_HUMAN  
39 1084.5 41.5 490 1 CPCJ\_HUMAN  
40 1083 41.4 491 1 CPHI\_CHICK  
41 1079.5 41.3 490 1 CPCR\_MBSAU  
42 1078.5 41.2 490 1 CPCH\_HUMAN  
43 1078.5 41.2 490 1 CPCT\_MOUSE  
44 1078.5 41.2 490 1 CPZ3\_MOUSE  
45 1072.5 41.0 490 1 CPCP\_MBSAU

## ALIGNMENTS

RESULT 1  
CPG1\_RABIT  
ID CPG1\_RABIT STANDARD; PRT; 494 AA.  
AC P24461;  
DT 01-MAR-1992 (Rel. 21, Created)  
DT 01-MAR-1992 (Rel. 21, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Cytochrome P450 2G1 (EC 1.14.14.1) (CYP11G1) (P450-NMB) (Olfactive).  
GN CYP2G1.  
OS Oryctolagus cuniculus (Rabbit).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
OX NCBI\_TaxID=9986;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91119408; PubMed=1703755;  
RA Ding X., Porter T.D., Peng H.M., Coon M.J.;  
RT "cDNA and derived amino acid sequence of rabbit nasal cytochrome  
P450NMB (P450IIG1), a unique isozyme possibly involved in  
olfaction."  
RL Arch. Biochem. Biophys. 285:120-125(1991).  
CC -!- FUNCTION: CYTOCHROMES P450 ARE A GROUP OF HEME-THIOLATE  
MONOOXYGENASES. THIS ISOZYME SEEMS TO BE IMPLICATED IN OLFACTION.  
CC -!- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +  
oxidized flavoprotein + H(2)O.  
CC -!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.  
CC -!- TISSUE SPECIFICITY: Olfactory epithelium.  
CC -!- SIMILARITY: Belongs to the cytochrome P450 family.  
PIR; S13907; B31944.  
HSP; P00179; I0T6.  
DR InterPro; IPR001128; Cytochrome P450.  
DR InterPro; IPR008067; EP450\_CYP2A.  
DR Pfam; PF00067; P450; 1.  
DR PRINTS; PR01684; EP450ICYP2A.  
DR PRINTS; PR00385; P450.  
DR PROSITE; PS00086; CYTOCHROME P450; 1.  
DR Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;  
KW Microsome; Endoplasmic reticulum; Olfaction.  
FT METAL 439 439 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).  
SQ SEQUENCE 494 AA; 56628 MW; 4DA83F2E1CD230BC CRC64;

Query Match 50.5%; Score 1319.5; DB 1; Length 494;  
Best Local Similarity 50.5%; Pred. No. 8.2e-85;  
Matches 248; Conservative 97; Mismatches 143; Indels 3; Gaps 3;  
QY 1 MEATGTWALLALALLLLTALSGTRAGHLPPGPTPLPLGNLQLRPGALYSGLMR 59  
Db 1 MELGGAFITFLALCSCLLLTAWKVRQVGRPLPGPTPIPLGNLQVTDATFOSFLK 60  
QY 60 LSKKGPVPTIYGLPWRPVVILVGOEAVREALGGQAEFSGRGTVMLEGTDFGHGVFFS 119  
Db 61 LREKGPVPTVYMGPRPVPVILCIGHEAVKEALVDRADEFSGRGELASVERNQCHGVALA 119  
QY 120 NGERWQRKFTMLALRDLGMRKEGELIQAEARCLVETFGTGTGPRPDPDLLLAQATS 179  
Db 120 NGERWRILRRFSLTILRDFGMGKRSEIERIQEAGYLLAEFRKTKGAPIDPTFFLSRVTS 179

RA Scott E.E., He Y.A., Wester M.R., White M.A., Chin C.C., Halpert J.R.,  
Johnson E.F., Stout C.D.; mammalian cytochrome P450 2B4 at 1.6-A  
"An open conformation."  
Proc. Natl. Acad. Sci. U.S.A. 100:13196-13201(2003).  
RL FUNCTION: CYTOCHROMES P450 ARE A GROUP OF HEME-THIOLATE  
MONOOXYGENASES. IN LIVER MICROSOMES, THIS ENZYME IS INVOLVED IN AN  
NADPH-DEPENDENT ELECTRON TRANSPORT PATHWAY. IT OXIDIZES A VARIETY  
OF STRUCTURALLY UNRELATED COMPOUNDS, INCLUDING STEROIDS, FATTY  
ACIDS, AND XENOBIOTICS. IN THE EPOXIDATION OF ARACHIDONIC ACID  
IT HAS A UNIQUE PREFERENCE FOR THE 5,6-OLEFIN.  
CC -!- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +  
oxidized flavoprotein + H(2)O.  
CC -!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.  
CC -!- INDUCTION: By phenobarbital.  
CC -!- POLYMORPHISM: Types B0 and B1 are probably allelic variants.  
CC -!- SIMILARITY: Belongs to the cytochrome P450 family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; M20856; AA65840.1; -;  
DR EMBL; M20857; AAA31224.1; -;  
DR PIR; A00179; O4RBPC.  
DR PDB; 1PO5; 07-OCT-03.  
DR InterPro; IPR001128; Cytochrome P450.  
DR InterPro; IPR008068; EP450\_CYP2B.  
DR Pfam; PF00067; P450; 1  
DR PRINTS; PR01685; EP450ICYP2B.  
DR PRINTS; PR00385; P450.  
DR PROSITE; PS00086; CYTOCHROME P450; 1.  
DR PROSITE; PS00086; Monooxygenase; Electron transport; Membrane; Iron;  
Heme; Microsome; Endoplasmic reticulum; Phosphorylation; Polymorphism;  
3D-structure. 128 128 PHOSPHORYLATION (BY PKA) (BY SIMILARITY).  
FT MOD\_RES 128 436 IRON (HEME AXIAL LIGAND).  
FT METAL 436 436 V -> I (IN BL).  
FT VARIANT 39 39 I -> V (IN BL).  
FT VARIANT 174 174 L -> I (IN BL).  
FT VARIANT 290 290 M -> L (IN BL).  
FT VARIANT 314 314 L -> M (IN BL).  
FT VARIANT 420 420 L -> M (IN BL).  
FT MUTAGEN 436 436 C->S: CONVERSION INTO AN NADPH OXIDASE  
WITH NEGLIGIBLE MONOOXYGENASE ACTIVITY.  
FT FT CONFLICT 91 91 Q -> E (IN REF. 3).  
FT CONFLICT 95 96 FS -> SF (IN REF. 3).  
FT CONFLICT 99 100 MISSING (IN REF. 3).  
FT CONFLICT 135 136 EG -> GY (IN REF. 3).  
FT CONFLICT 193 193 P -> K (IN REF. 3).  
FT CONFLICT 221 221 P -> S (IN REF. 3).  
FT CONFLICT 303 303 T -> A (IN REF. 3).  
FT CONFLICT 461 465 SPVPP -> GNLSL (IN REF. 3).  
FT SEQUENCE 491 AA; 55713 MW; 0DB943C6CDCF690B CRC64;  
Query Match 49.7%; Score 1300; DB 1; Length 491;  
Best Local Similarity 50.1%; Pred. No. 1.9e-83;  
Matches 244; Conservative 101; Mismatches 138; Indels 4; Gaps 3;  
QY 7 WALLLALALL--LILLTALSGTRARGHLPPGTPPLPLGLNLLQRLPGLALYSGLMRLSKY 64  
Db 3 FSLILLALLAFGLALLLLFRGHPRAGHLPPGTPPLVGLNLLQMDRKLGLRLREKY 62  
QY 65 GPVFTIYGLPWRPVVVLVQEAVERALGGQAEBSGRGTVMLEGTGDFGCVFSGNBERW 124  
Db 63 GDVFTVYLGSRPVPVVLCTGDAREALVDQAEAFSGRGKIAVWDPIFGYGVIFANGERW 121  
QY 125 RQLRKFTMLALRDLGCMGRKEGELIQAEARCLVETFOGTGPRPDPSSLALLAQATSNVVC 184  
Db 122 RAURRFLATMRDFGMRKRSVEERIQEARCLVETFOGTGPRPDPSSLALLAQATSNVVC 181

RA  
RA Johnson E.F., Stout C.D.;  
RT "An open conformation";  
RL Proc. Natl. Acad. Sci. U.S.A. 100:13196-13201(2003).  
CC -!- FUNCTION: CYTOCHROMES P450 ARE A GROUP OF HEME-THIOLATE  
MONOOXYGENASES. IN LIVER MICROSOMES, THIS ENZYME IS INVOLVED IN AN  
NADPH-DEPENDENT ELECTRON TRANSPORT PATHWAY. IT OXIDIZES A VARIETY  
OF STRUCTURALLY UNRELATED COMPOUNDS, INCLUDING STEROIDS, FATTY  
ACIDS, AND XENOBIOTICS. IN THE EPOXIDATION OF ARACHIDONIC ACID  
IT HAS A UNIQUE PREFERENCE FOR THE 5,6-OLEFIN.  
CC -!- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +  
oxidized flavoprotein + H(2)O.  
CC -!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.  
CC -!- INDUCTION: By phenobarbital.  
CC -!- POLYMORPHISM: Types B0 and B1 are probably allelic variants.  
CC -!- SIMILARITY: Belongs to the cytochrome P450 family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; M20856; AA65840.1; -;  
DR EMBL; M20857; AAA31224.1; -;  
DR PIR; A00179; O4RBPC.  
DR PDB; 1PO5; 07-OCT-03.  
DR InterPro; IPR001128; Cytochrome P450.  
DR InterPro; IPR008068; EP450\_CYP2B.  
DR Pfam; PF00067; P450; 1  
DR PRINTS; PR01685; EP450ICYP2B.  
DR PRINTS; PR00385; P450.  
DR PROSITE; PS00086; CYTOCHROME P450; 1.  
DR PROSITE; PS00086; Monooxygenase; Electron transport; Membrane; Iron;  
Heme; Microsome; Endoplasmic reticulum; Phosphorylation; Polymorphism;  
3D-structure. 128 128 PHOSPHORYLATION (BY PKA) (BY SIMILARITY).  
FT MOD\_RES 128 436 IRON (HEME AXIAL LIGAND).  
FT METAL 436 436 V -> I (IN BL).  
FT VARIANT 39 39 I -> V (IN BL).  
FT VARIANT 174 174 L -> I (IN BL).  
FT VARIANT 290 290 M -> L (IN BL).  
FT VARIANT 314 314 L -> M (IN BL).  
FT VARIANT 420 420 L -> M (IN BL).  
FT MUTAGEN 436 436 C->S: CONVERSION INTO AN NADPH OXIDASE  
WITH NEGLIGIBLE MONOOXYGENASE ACTIVITY.  
FT FT CONFLICT 91 91 Q -> E (IN REF. 3).  
FT CONFLICT 95 96 FS -> SF (IN REF. 3).  
FT CONFLICT 99 100 MISSING (IN REF. 3).  
FT CONFLICT 135 136 EG -> GY (IN REF. 3).  
FT CONFLICT 193 193 P -> K (IN REF. 3).  
FT CONFLICT 221 221 P -> S (IN REF. 3).  
FT CONFLICT 303 303 T -> A (IN REF. 3).  
FT CONFLICT 461 465 SPVPP -> GNLSL (IN REF. 3).  
FT SEQUENCE 491 AA; 55713 MW; 0DB943C6CDCF690B CRC64;  
Query Match 49.7%; Score 1300; DB 1; Length 491;  
Best Local Similarity 50.1%; Pred. No. 1.9e-83;  
Matches 244; Conservative 101; Mismatches 138; Indels 4; Gaps 3;  
QY 7 WALLLALALL--LILLTALSGTRARGHLPPGTPPLPLGLNLLQRLPGLALYSGLMRLSKY 64  
Db 3 FSLILLALLAFGLALLLLFRGHPRAGHLPPGTPPLVGLNLLQMDRKLGLRLREKY 62  
QY 65 GPVFTIYGLPWRPVVVLVQEAVERALGGQAEBSGRGTVMLEGTGDFGCVFSGNBERW 124  
Db 63 GDVFTVYLGSRPVPVVLCTGDAREALVDQAEAFSGRGKIAVWDPIFGYGVIFANGERW 121  
QY 125 RQLRKFTMLALRDLGCMGRKEGELIQAEARCLVETFOGTGPRPDPSSLALLAQATSNVVC 184  
Db 122 RAURRFLATMRDFGMRKRSVEERIQEARCLVETFOGTGPRPDPSSLALLAQATSNVVC 181

QY 180 NVVCSLLFLGRFVSYEDKEFOAVVRAAGTLLGVSSQGGQTYEMFSWFLRPLPGRPHQLH 239  
Db 180 NVISVVGSGFDYEDKQFLSLLRWINESEFTEMSTPWAQLYDMYSGVMQYLPGRHRIY 239  
QY 240 HVSTLAFTVRQVQOQHNDASGARDLVDAFLLKMAQEQNPCTETNKMLMTVYL 299  
Db 240 LIEELKDFAARVKVNEASLDQNP-RDFIDCFLLKHQDKNNPHTEFNKLVLTNL 298  
QY 300 LPAGTMTVSTVGYTLILLMKYPHYQVKKVRELNRELGAQPSLGDRTPLPYTDVHLR 359  
Db 299 FFAGTETVSSRLRYGFLIMKHPEVQTKIYEINQVGHRIPSVDDRVKPFDDVHLR 358  
QY 360 AQRLLALVPMGIPRLMTTRPRGYTLPGTEVFPLGLSILHDPNIFKHPEEFNDRFLD 419  
Db 359 IQLRTDIPVPMGVPHNVIRDTFRGYLLPKGTDVFLGLSVLKDPKYFCHPDDFYQHELD 418  
QY 420 ADGRFRKHEAFPLFSLGRKRVCLGELAKALELFTTILQAFSLSPCPDPLSLKPTVS 479  
Db 419 EQGREKKEAFVFPSSGRKICLGEAMRMELFLYFTSLQNFSLHPLVPPVNVNIDITPKIS 478  
QY 480 GLFNIPPAFQL 490  
Db 479 GFGNIPPTVEL 489

RESULT 2  
CPB4\_RABIT STANDARD; PRT; 491 AA.  
AC P00178; P00177;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 15-WAR-2004 (Rel. 43, Last annotation update)  
DE Cytochrome P450 2B4 (EC 1.14.14.1) (CYP11B4) (P450-LM2) (Isozyme 2)  
DE (P450 types B0 and B1).  
GN CYP2B4.  
OS Oryctolagus cuniculus (Rabbit).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
OX NCBI\_TaxID=9986;  
[1]  
RN SEQUENCE FROM N.A.  
RX MEDLINE=88094290; PubMed=2826996;  
RA Gasser R., Negishi M., Philpot R.M.;  
RT "Primary structures of multiple forms of cytochrome P-450 isozyme 2  
derived from rabbit pulmonary and hepatic cDNAs";  
RL Mol. Pharmacol. 33:22-30(1988).  
[2]  
RN SEQUENCE.  
RX MEDLINE=84042509; PubMed=6579541;  
RA Tarr G.E., Black S.D., Fujita V.S., Coon M.J.;  
RT "Complete amino acid sequence and predicted membrane topology of  
phenobarbital-induced cytochrome P-450 (isozyme 2) from rabbit liver  
microsomes";  
RL Proc. Natl. Acad. Sci. U.S.A. 80:6552-6556(1983).  
[3]  
RN SEQUENCE.  
RX MEDLINE=83160983; PubMed=6833251;  
RA Heinemann F.S., Ozols J.;  
RT "The complete amino acid sequence of rabbit phenobarbital-induced  
liver microsomal cytochrome P-450";  
RL J. Biol. Chem. 258:4195-4201(1983).  
[4]  
RN MUTAGENESIS OF CYS-436.  
RX MEDLINE=22222627; PubMed=12237221;  
RA Vatsis K.P., Peng H.-M., Coon M.J.;  
RT "Replacement of active-site cysteine-436 by serine converts cytochrome  
P450 2B4 into an NADPH oxidase with negligible monooxygenase  
activity";  
RL J. Inorg. Biochem. 91:542-553(2002).  
[5]  
RN X-RAY CRYSTALLOGRAPHY (1.6 ANGSTROMS).  
RP MEDLINE=22975166; PubMed=14563924;



KW	Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
KW	Microsome; Endoplasmic reticulum; Phosphorylation.
FT	MOD_RES 128 128
FT	METAL 436 436
FT	CONFLICT 292 292
FT	CONFLICT 321 321
FT	CONFLICT 322 322
FT	CONFLICT 438 438
FT	CONFLICT 444 444
FT	CONFLICT 473 473
FT	CONFLICT 476 476
FT	CONFLICT 491 AA; 55932 MW; 00CB68937FDD44BC CRC64;
SEQUENCE	491 AA; 55932 MW; 00CB68937FDD44BC CRC64;
Query Match 49.6%; Score 1296; DB 1; Length 491;	
Best Local Similarity 50.4%; Pred. No. 3.5e-83;	
Matches 243; Conservative 97; Mismatches 140; Indels 2; Gaps 2;	
QY	9 LLLALLLLLTALSGTRARGHLPPGPTPLLLGNLLQLRPGALYSGLMRLSKYGPVF 68
DB	7 LLLALLVGLLLVVRGHPKSRGNFPFGRPLPLLLGNLLQLDRGGLNLFMQREKYGDFV 66
QY	69 TIVLGPMPVVVLVGOEAVREALGQAEFSGRGTVAMLEGTFDGHGVFFSNGERWRQLR 128
DB	67 TVHLGP-RPVVMLCGTDITKEALVQAEDFSGRGTIAVIEPIFKEYGVIFANGERWKAIR 125
QY	129 KFTMLALRDLGMRGKEGELIOAEARCLVETFCGTGREGPPDPSLLLAQATSNVVCSLIFG 188
DB	126 RPSLATMRDFGMGKRSVEERIOEAAQCLVEELRKSQGAFLDPTFLPQCITANIICSVIFG 185
QY	189 LRFPSVEDKEFOAVVRAAGTGLGVSSQGGTQVEMFSWFLRPLPGPHKQLLHHVSTLAAFT 248
DB	186 ERFDTYDROFLLELFYFTLSLLSFSQVFEFFSGFLKYPFGAHRQISKNLOELDIYI 245
QY	249 VRQVOGHQNDASGPARDLVDAFLKMAQEONPGTEFTNKNMLMTVIYLLFAGTMTVS 308
DB	246 GHIVEKHRTLDPASAP-RDFIDTYLLRMEKESKNHTEPHENMLISLSLFFAGTETGS 304
QY	309 TTIVGYTLMLLMKYPHQKWRRELNRELGAQAPSLGDRTRLPYTDVAVHEAQRLLALVP 368
DB	305 TTRYGYTLMLKYPHTEKVKQEKIDQVIGSHRPPSLDDRTKMPYTDVAVHEIQRFADLAP 364
QY	369 MGIPRLMTTRRGYTLPGQTEVFPLGSLIHDPNI EKHPBEFNDRFLDADGRPKHE 428
DB	365 IGLPHRVTKDTMFRGYLLPKNTEVPILSSALHDPOYFDHPDTFNEHEFLDADGTIKXE 424
QY	429 AFLPFSGLGRVCLGEGAKAELEFFTTILQAFLSPSPCPDPTLSLKPTVSGFLNPAPF 488
DB	425 AFWPFSGTGRICLGEIGARNELFFTTILQNFVSSSHAPKIDILTPKESGIKIPPTY 484
QY	489 QL 490
DB	485 QI 486
RESULT 4	
ID	CPB5_RABIT STANDARD; PRT; 491 AA.
AC	P12789;
DT	01-OCT-1989 (Rel. 12, Created)
DT	01-OCT-1989 (Rel. 12, Last sequence update)
DT	28-FEB-2003 (Rel. 41, Last annotation update)
DE	Cytochrome P450 2B5 (EC 1.14.14.1) (CYF11B5) (P450 type B2) (P450 form HP1).
DE	CYP2B5.
GN	Oryctolagus cuniculus (Rabbit).
OS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX	NCBI_TaxID=9986;
RN	[1]
RP	SEQUENCE FROM N.A.
RX	MEDLINE=88094290; PubMed=2826996;
RA	Gaesser R., Negishi M., Philpot R.M.;
RT	"Primary structures of multiple forms of cytochrome P-450 isozyme 2

derived from rabbit pulmonary and hepatic cDNAs.";

Mol. Pharmacol. 33:22-30(1988).

[2]

SEQUENCE OF 1-56 FROM N.A.

MEDLINE=95383381; PubMed=7654758;

Lehner M., Schulze J., Petzold A., Bernhardt R., Hlavica P.;

"Rabbit liver cytochrome P-450 2B5: high-level expression of the

full-length protein in Escherichia coli, purification, and catalytic

activity.";

Biochim. Biophys. Acta 1245:107-115(1995).

FUNCTION: Cytochromes P450 are a group of heme-thiolate

monooxygenases. In liver microsomes, this enzyme is involved in an

NADPH-dependent electron transport pathway. It oxidizes a variety

of structurally unrelated compounds, including steroids, fatty

acids, and xenobiotics.

CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +

oxidized flavoprotein + H(2)O.

SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.

INDUCTION: By phenobarbital.

SIMILARITY: Belongs to the cytochrome P450 family.

SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/) or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; M20855; AAA31223.1; -;

EMBL; S78830; AAB35177.1; -;

PIR; S31278; S31278.

HSSP; P00179; 1DT6.

InterPro; IPR001128; Cytochrome P450.

DR	InterPro; IPR008068; EP450\_CYP25.
DR	Pfam; PF00067; P450; 1.
DR	PRINTS; PR01685; EP450ICYP2B.
DR	PRINTS; PR00385; P450.
DR	PROSITE; PS00086; CYTOCHROME P450; 1.
DR	Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
KW	Microsome; Endoplasmic reticulum; Phosphorylation.
FT	MOD\_RES 128 128
FT	METAL 436 436
FT	SEQUENCE 491 AA; 55769 MW; ABE2B978B8408DCF CRC64;
SEQUENCE	491 AA; 55769 MW; ABE2B978B8408DCF CRC64;
Query Match 49.6%; Score 1296; DB 1; Length 491;	
Best Local Similarity 50.3%; Pred. No. 3.5e-83;	
Matches 245; Conservative 97; Mismatches 141; Indels 3; Gaps 3;	
QY	7 WALLALALL--LLLTALSGTRARGHLPPGPTPLLLGNLLQLRPGALYSGLMRLSKY 64
DB	3 FSLILLALFAGLLLLFRGHPKAHGLPPGPPPLVGLNLLQMDRKGLRLSRFLREKY 62
QY	65 GPVFTIYLGWRPVVVLVGOEAVREALGQAEFSGRGTVAMLEGTFDGHGVFSNGERM 124
DB	63 GDVFTVVLGS-RPVVVLGCTDAIREALVDQAEAFSGRGTVAVDPIFGQGVFPFANGHW 121
QY	125 RQLRKETMLALRDLGMRGKEGELIOAEARCLVETFCGTGREGPPDPSLLLAQATSNVVC 184
DB	122 RALRRFSLATMRDFGMGKRSVEERIOEAEARCLVEELRKSQGAFLDPTFLPQCITANIIC 181
QY	185 LFLGLAFSVDKBFQAVVRAAGTGLGVSSQGGTQVEMFSWFLRPLPGPHKQLLHHVSTL 244
DB	182 IVFGKRFYKDPVFLRLDLFFQSFSLISSFSQVLELFPGLKHFPGTHQRYRNQEI 241
QY	245 AATVTRQVOOHQGNLDASGPARDLVDAFLKMAQEONPGTEFTNKNMLMTVIYLLFAGT 304
DB	242 NTFIGQTVEXHRATLDPNSP-RDFIDVYLLRMKDKSDPSSEFHHRLIITVTLTFAGT 300
QY	305 MTVSTTVGYTLLLMKYPHQKWRRELNRELGAQAPSLGDRTRLPYTDVAVHEAQRLL 364
DB	301 ETTSTTLRYGFLMLKYPHTEKVKQEKIEQVIGSHRPPALDDRAKMPYTDVAVHEIQRLG 360



DE hydroxylase) (P450-16-alpha) (Clone PF3/46).  
GN CYP2B10 OR CYP2B-10.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=89118235; PubMed=3219345;  
RA Noshiro M., Lakso M., Kawajiri K., Negishi M.,  
RT "Rip locus: regulation of female-specific isozyme (I-P-450(16 alpha)  
RT of testosterone 16 alpha-hydroxylase in mouse liver, chromosome  
RT localization, and cloning of P-450 cDNA.";  
RL Biochemistry 27:6434-6443(1988).  
CC -!- FUNCTION: Cytochromes P450 are a group of heme-thiolate  
CC monooxygenases. In liver microsomes, this enzyme is involved in an  
CC NADPH-dependent electron transport pathway. It oxidizes a variety  
CC of structurally unrelated compounds, including steroids, fatty  
CC acids, and xenobiotics.  
CC -!- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +  
CC oxidized flavoprotein + H(2)O.  
CC -!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.  
CC -!- SIMILARITY: Belongs to the cytochrome P450 family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC  
CC EMBL; M21856; AAA40425.1; -.  
CC PIR; B31047; B31047.  
CC HSP; P00179; 1DT6.  
CC MGD; MGI:88598; Cyp2b10.  
CC InterPro; IPR001128; Cytochrome\_P450.  
CC InterPro; IPR008068; EP450\_CYP2B.  
CC Pfam; PF00067; P450; 1.  
CC PRINTS; PR01685; EP450ICYP2B.  
CC PROSITE; PR00385; P450.  
CC PROSITE; PS00086; CYTOCHROME\_P450; FALSE NEG.  
CC Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;  
KW Microsome; Endoplasmic reticulum; Phosphorylation.  
FT MOD RES 128 PHOSPHORYLATION (BY PKA) (BY SIMILARITY).  
FT METAL 445 IRON (HEME AXIAL LIGAND).  
FT SEQUENCE 500 AA; 56743 MW; F660A00D8D0FBA94 CRC64;  
CC  
CC Query Match 49.0%; Score 1280.5; DB 1; Length 500;  
CC Best Local Similarity 49.7%; Pred. No. 4.4e-82;  
CC Matches 244; Conservative 99; Mismatches 137; Indels 11; Gaps 3;  
CC  
CC 9 LLLALLLLTLLALSGTRAGHLPPOPTPLPLGLNLLQLRPGALYSGMLSKKYGVPF 68  
CC 7 LLLALLVGFLLLLARHPKSGRGNFPQPRPLPLGLNLLQMDRGGLLSLIQLREKYGDVF 66  
CC  
CC 69 TIYLGPMRPVVLVQBAVREALGGQAEFSGTGMLEGTDPDGHGVFNSGERWQLR 128  
CC 67 TVHLGP-PPVVMVLCGTDTREALVQAEAFSGRTVAVVEPTFKEYGVIFANGEMKTLR 125  
CC  
CC 129 KFTMLALDLMGKREGEELIOAEARCLVETFOGTEGRPDPSLLLAQATSNVVCSLFPG 188  
CC 126 RFLSLATWRDPMGKRSVEERIQEAQCLVEELRSQGPAPDPTFLFOCIATNVCISLVFG 185  
CC  
CC 189 LRFSYEDKEFOAVRAAGGTLGLVSSGGQTYEFSWFLRPLPGHPKQLLHHVSTLAFT 248  
CC 186 EREYTDQFRLLELLEFYQVFLSLSSFSQQMFLSGFLKYPFGAHRQISKNLQELLDYI 245  
CC 249 VRQVQHQGNLDASGPARDLVDAFLKMAQEQNPNGTEFTNKNMLMTVILLFAGTIVTS 308  
CC 246 GHSVERHKATLDPSVP-RDFTIDVLLRMEKEKSNQNAEFHQNLMMVSLSLFFVGTETSS 304  
CC 309 TTVGTYTLLLLMKYHVKVQWREELNRELGAQAPSLGDRTRLPTVDVLAHQRLALVP 368

Db 305 TLLHYGFLMLKYPHVTEKQKEIDQVIGSHRLPTLDDRTKMPYSDAVIHEIORFSDLIP 364  
Qy 369 MGIPRTLMRTTRPRGYLYLPGQTEVFPPLGLSILHDPNIFKHEEENPDRLDADGRFKHE 428  
Db 365 IGVHRTVKTDLFRGYLLPKNTTEVYPISSLALHDPQYFEQPDSENPDPFLDANGALKKSE 424  
Qy 429 AFLRP-----SLGKRVCLGEGAKAEFLPFTTILQAFSLAESPCPPDTSLSLKPTVS 479  
Db 425 AFLPSTQGIIFDQKSVGKRICLGESIAESELFLFTSLQNFVASHVAPKXIDITPKES 484  
Qy 480 GLFNIPPAFQL 490  
Db 485 GIGKIPPTYQI 495  
CC  
CC RESULT 7  
CC ID\_CPB CANFA STANDARD; PRT; 494 AA.  
CC AC P24460;  
CC DT 01-MAR-1992 (Rel. 21, Created)  
CC DT 01-MAR-1992 (Rel. 21, Last sequence update)  
CC DT 28-FEB-2003 (Rel. 41, Last annotation update)  
CC DE Cytochrome P450 2B11 (EC 1.14.14.1) (CYP11B1) (P450 PBD-2).  
CC GN CYP2B11.  
CC OS Canis familiaris (Dog).  
CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
CC OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
CC OX NCBI\_TaxID=9615;  
CC RN [1]  
CC RP SEQUENCE FROM N.A.  
CC RC STRAIN=Beagle; TISSUE=Liver;  
CC RX MEDLINE=9034348; PubMed=2116765;  
CC RA Graves P.E., Elhag G.A., Ciaccio P.J., Bourque D.P., Halpert J.R.;  
CC RT "cDNA and deduced amino acid sequences of a dog hepatic cytochrome  
CC RT P45011B responsible for the metabolism of 2',4',4',5',5'-  
CC RT hexachlorobiphenyl.";  
CC RL Arch. Biochem. Biophys. 281:106-115(1990).  
CC -!- FUNCTION: CYTOCHROMES P450 ARE A GROUP OF HEME-THIOLATE  
CC MONOOXYGENASES. IN LIVER MICROSOMES, THIS ENZYME IS INVOLVED IN AN  
CC NADPH-DEPENDENT ELECTRON TRANSPORT PATHWAY. THIS ISOZYME SEEMS  
CC RESPONSIBLE FOR METABOLISM OF 2',4',4',5',5'-HEXACHLOROBIPHENYL.  
CC -!- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +  
CC oxidized flavoprotein + H(2)O.  
CC -!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.  
CC -!- INDUCTION: By phenobarbital.  
CC -!- SIMILARITY: Belongs to the cytochrome P450 family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC  
CC EMBL; M92447; AAA30881.1; -.  
CC PIR; S11305; S11305.  
CC HSP; P00179; 1DT6.  
CC InterPro; IPR001128; Cytochrome\_P450.  
CC InterPro; IPR008068; EP450\_CYP2B.  
CC Pfam; PF00067; P450; 1.  
CC PRINTS; PR01685; EP450ICYP2B.  
CC PROSITE; PR00385; P450.  
CC PROSITE; PS00086; CYTOCHROME\_P450; 1.  
CC Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;  
KW Microsome; Endoplasmic reticulum; Phosphorylation.  
FT MOD RES 128 PHOSPHORYLATION (BY PKA) (BY SIMILARITY).  
FT METAL 436 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).  
FT SEQUENCE 494 AA; 56266 MW; 8E9EA639B4B5E4D3 CRC64;  
CC  
CC Query Match 48.3%; Score 1262; DB 1; Length 494;  
CC Best Local Similarity 49.2%; Pred. No. 8.4e-81;



Matches 237; Conservative 99; Mismatches 144; Indels 2; Gaps 2;

QY 9 LLLALLLLTLLALSGTRARGHLPPTPLPILGNLLQLRPGALYSGLMRLSKYGVVF 68  
 DB 7 LLLALLTLLLMARGHKAHYLPGRPRPILGNFLQMDRKGLKSLFLRLOEKYGVVF 66  
 QY 69 TIYGLPWPVVVLVGOEAVREALGQAEFSGRTVAMLEGTDFGCHGVFFNGERWRQLR 128  
 DB 67 TVYGLP-RTVWLCGIDAIREALVDNAEAFSGRKIAVVEPFGYGVVFANGERWKTLR 125  
 QY 129 KFTMLALRDLGMKREGEELIQAERCLIVETFOGTGEPFDPDPSLLLAQATSNVVCSSLFG 188  
 DB 126 RFLSLATMRDGMKRSVEERIQAERCLIVETFOGTGEPFDPDPSLLLAQATSNVVCSSLFG 185  
 QY 189 LRFSEYDEKFEQAVVRAAGTLLGVSSGGQGVYEMFSLRPLGPHKOLLHVVSTLAFT 248  
 DB 186 KRFYKQPEFLRMNLFVFSALISSFSQMFELFHSFLKYPFGTHRQVYNNLQEIKAFI 245  
 QY 249 VRQVQHQGNLDASGPARDLVDAFLKMAQEQNGTEFTNKNMLMTVIVLLFAGTMTVS 308  
 DB 246 ARMYKEHRETLDPSAP-RDPIDAYLIRMDKEAEPSEFHHNLDLTALSFFAGTMTVS 304  
 QY 309 TTVCYGTLLLMKYPHVQKWRRELNLRELGAQAPSLGDRTPYTDVAVLHEAQRLLALVP 368  
 DB 305 TTYRYGFLMLKYPHIAERIYKEIDQVIGPHRLPSLDRAKQPYTDVAVIHEIQRFGLLP 364  
 QY 369 MGIPRTLMRTTRFRGYTLPGCTEVEPLLSGLTLHDPNIFKHPPEENPDRFLDADGRFKHE 428  
 DB 365 IGVPHMYTKICFRGIYIPKTEVFPILHNSALNDPHYFEPKPDVFNPDHFLDANGALKKNE 424  
 QY 429 AFIPFSGKRCVLCGEGAKAEELFFFTTILQAFSLSPCPDPTLSLKPTVSGLFNIPPAF 488  
 DB 425 AFIPFSGKRCVLCGEGAKAEELFFFTTILQAFSLSPCPDPTLSLKPTVSGLFNIPPAF 484  
 QY 489 QL 490  
 DB 485 QI 486

RESULT 8 8  
 CPB9 MOUSE  
 ID \_CPB9\_MOUSE STANDARD; PRT; 491 AA.  
 AC P12790;  
 DT 01-OCT-1989 (Rel. 12, Created)  
 DT 01-OCT-1989 (Rel. 12, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Cytochrome P450 2B9 (EC 1.14.14.1) (CYP1B9) (Testosterone 16-alpha hydroxylase) (P450-16-alpha) (Clone PF26).  
 GN CYP2B9 OR CYP2B-9 OR RIP.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]\_TaxID=10090;  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver;  
 RX MEDLINE=89118235; PubMed=3219345;  
 RA Noshiro M., Lakso M., Kawajiri K., Negishi M.;  
 RT "Rip locus: regulation of female-specific isozyme (I-P-450(16 alpha) of testosterone 16 alpha-hydroxylase in mouse liver, chromosome localization, and cloning of P-450 cDNA.";  
 RL Biochemistry 27:6434-6443 (1988).  
 CC -!- FUNCTION: Cytochromes P450 are a group of heme-thiolate monooxygenases. In liver microsomes, this enzyme is involved in an NADPH-dependent electron transport pathway. It oxidizes a variety of structurally unrelated compounds, including steroids, fatty acids, and xenobiotics.  
 CC -!- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH + oxidized flavoprotein + H(2)O.  
 CC -!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.  
 CC -!- SIMILARITY: Belongs to the cytochrome P450 family.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).

EMBL; M21855; AAA40424.1; -;  
 PIR; A31047; A31047.  
 DR HSSP; P00179; 1DT6.  
 DR MGD; MGI:88600; Cyp2b9.  
 DR InterPro; IPR001128; Cytochrome P450.  
 DR InterPro; IPR008068; EP450\_CYP2B.  
 DR Pfam; PF00067; P450; 1.  
 DR PRINTS; PR01685; EP450CYP2B.  
 DR PRINTS; PR00385; P450.  
 DR PROSITE; PS00086; CYTOCHROME P450; 1.  
 KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;  
 KW Microneme; Endoplasmic reticulum; Phosphorylation.  
 FT MOD\_RES 128 128 PHOSPHORYLATION (BY PKA) (BY SIMILARITY).  
 FT METAL 436 436 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).  
 SQ SEQUENCE 491 AA; 55760 MW; E626F61F136CF697 CRC64;

Query Match 47.68; Score 1244; DB 1; Length 491;  
 Best Local Similarity 48.5%; Pred. No. 1.5e-79;  
 Matches 234; Conservative 100; Mismatches 146; Indels 2; Gaps 2;

QY 9 LLLALLLLTLLALSGTRARGHLPPTPLPILGNLLQLRPGALYSGLMRLSKYGVVF 68  
 DB 7 LLLALLTLLLMARGHKAHYLPGRPRPILGNFLQMDRKGLKSLFLRLOEKYGVVF 66  
 QY 69 TIYGLPWPVVVLVGOEAVREALGQAEFSGRTVAMLEGTDFGCHGVFFNGERWRQLR 128  
 DB 67 TVYGLP-RPVVVLGCTQITREALVDNAEAFSGRTVAMLEGTDFGCHGVFFNGERWRQLR 125  
 QY 129 KFTMLALRDLGMKREGEELIQAERCLIVETFOGTGEPFDPDPSLLLAQATSNVVCSSLFG 188  
 DB 126 RFLSLATMRDGMKRSVEERIQAERCLIVETFOGTGEPFDPDPSLLLAQATSNVVCSSLFG 185  
 QY 189 LRFSEYDEKFEQAVVRAAGTLLGVSSGGQGVYEMFSLRPLGPHKOLLHVVSTLAFT 248  
 DB 186 ERFDTYDDQFLHLLNLMYKIFLSLSSFGQMFELFHSFLKYPFGTHRQVYNNLQEIKAFI 245  
 QY 249 VRQVQHQGNLDASGPARDLVDAFLKMAQEQNGTEFTNKNMLMTVIVLLFAGTMTVS 308  
 DB 246 ARMYKEHRETLDPSAP-RDYIDTYLLRMEKESNHNTEFHHNLDLTALSFFAGTMTVS 304  
 QY 309 TTVCYGTLLLMKYPHVQKWRRELNLRELGAQAPSLGDRTPYTDVAVLHEAQRLLALVP 368  
 DB 305 ATLYGVLLMLKYPHVTKEVQKBIDQVIGSHRLPTLDDRTKMPYTDVAVIHEIQRFGLLP 364  
 QY 369 MGIPRTLMRTTRFRGYTLPGCTEVEPLLSGLTLHDPNIFKHPPEENPDRFLDADGRFKHE 428  
 DB 365 IGVPHMYTKICFRGIYIPKTEVFPILHNSALNDPHYFEPKPDVFNPDHFLDANGALKKNE 424  
 QY 429 AFIPFSGKRCVLCGEGAKAEELFFFTTILQAFSLSPCPDPTLSLKPTVSGLFNIPPAF 488  
 DB 425 AFIPFSGKRCVLCGEGAKAEELFFFTTILQAFSLSPCPDPTLSLKPTVSGLFNIPPAF 484  
 QY 489 QL 490  
 DB 485 QI 486

RESULT 9

CPGL RAT

ID CPGL RAT STANDARD; PRT; 494 AA.

AC P10610; Q64589;

DT 01-JUL-1989 (Rel. 11, Created)

DT 01-JUL-1989 (Rel. 11, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Cytochrome P450 2G1 (EC 1.14.14.1) (CYP1G1) (P450-OLFI) (Olfactive).

GN CYP2G1 OR CYP2G-1.

OS	Rattus norvegicus (Rat).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.	
OX	NCBI_TaxID=10116;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RP	MEDLINE=90153922; PubMed=2406242;	
RA	Nef P., Larabee T.M., Kagimoto K., Meyer U.A.;	
RA	"Olfactory-specific cytochrome P-450 (P-450olf1; IIG1). Gene	
RT	structure and developmental regulation.";	
RT	J. Biol. Chem. 265:2903-2907(1990).	
RL	[2]	
RN	SEQUENCE OF 2-494 FROM N.A.	
RP	MEDLINE=89214086; PubMed=2708343;	
RA	Nef P., Heldman J., Lazard D., Margalit T., Jaye M., Hanukoglu I.,	
RA	Lancet D.;	
RT	"Olfactory-specific cytochrome P-450. cDNA cloning of a novel	
RT	neuroepithelial enzyme possibly involved in chemoreception.";	
RL	J. Biol. Chem. 264:6780-6785(1989).	
CC	- FUNCTION; CYTOCHROMES P450 ARE A GROUP OF HEME-THIOLATE	
CC	MONOOXYGENASES. THIS ISOZYME SEEMS TO BE IMPLICATED IN OLFACTION.	
CC	- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +	
CC	oxidized flavoprotein + H(2)O.	
CC	- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.	
CC	- TISSUE SPECIFICITY: Olfactory epithelium.	
CC	- SIMILARITY: Belongs to the cytochrome P450 family.	
CC	-----	
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration	
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -	
CC	the European Bioinformatics Institute. There are no restrictions on its	
CC	use by non-profit institutions as long as its content is in no way	
CC	modified and this statement is not removed. Usage by and for commercial	
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>	
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).	
CC	-----	
DR	EMBL; M31931; AAA41069.1; -	
DR	EMBL; M31923; AAA41069.1; JOINED.	
DR	EMBL; M31924; AAA41069.1; JOINED.	
DR	EMBL; M31925; AAA41069.1; JOINED.	
DR	EMBL; M31926; AAA41069.1; JOINED.	
DR	EMBL; M31927; AAA41069.1; JOINED.	
DR	EMBL; M31928; AAA41069.1; JOINED.	
DR	EMBL; M31929; AAA41069.1; JOINED.	
DR	EMBL; M31930; AAA41070.2; JOINED.	
DR	EMBL; M33296; AAA41070.2; -	
DR	PIR; A35551; A35551.	
DR	HSP; P00179; 1DT6.	
DR	InterPro; IPR001128; Cytochrome P450.	
DR	InterPro; IPR008067; EP450_CYP2A.	
DR	Pfam; PF00067; p450; 1.	
DR	PRINTS; PR01684; EP450ICYP2A.	
DR	PRINTS; PR00385; P450.	
DR	PROSITE; PS00086; CYTOCHROME P450; 1.	
KW	Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;	
KW	Microsome; Endoplasmic reticulum; Olfaction.	
FT	METAL 439 439 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).	
SQ	SEQUENCE 494 AA; 56780 MW; 8A06E256A9A48702 CRC64;	
	Query Match 47.6%; Score 1243.5; DB 1; Length 494;	
	Best Local Similarity 47.7%; Pred. No. 1.6e-79;	
	Matches 234; Conservative 101; Mismatches 153; Indels 3; Gaps 3;	
QY	1 MEATGTWALLALAL-LLLTALSGTRAGHLPFGTPTPLGNLLQLRPGALYSGLMR 59	
Db	1 MALGGAFSPMTCLSLILIAWKRTSRGGKLPFGTPTPIPLGNLLQVRIDATFOSFLK 60	
QY	60 LSKKYGPVTIYLGPMRPVVVLGQAVREALGQAEFGSGRGTVAMLEGTFCGHGVFFS 119	
Db	61 LQKYGVSFTVYFGR-PPVILCGHEAVKCALVDQDDFGRGEMPTLEKNFGYGLALS 119	
QY	120 NGRWRQLRFTMLALDLGMKRGEBELIQAEARCLIVETFCQTGEGRPFPFSLLAQATS 179	
Db	120 NGRWKILRFLSLVLRNFGMGKSEIERIQEEAGYLLELHKVKGAPIDPTFTYLSRTVS 179	

QY	180 NVVCSLLFGLRFSYEDKEFOAVVRAAGTLLGVSSQGGQTYEMFSWFLRPLPGPHKQLLH 239	
Db	180 NVICSVVFGKRFYEDQRFSLMKMINESFVEMSPWAQLYDMYGVVIQYFGRHRLYN 239	
QY	240 HVSTLAFTVRQVQOQHOGLNDASGPARDLVDAFLLKMAQEQNPGTGTETNKMLMTVIYL 299	
Db	240 LIEELKOPIASRVKINEASFDPSNP-RDFIDCFLLKMYQDKSDPHSEENKLNVLTLNL 298	
QY	300 LFAGTMVSTVGTGTTTTLLLMKYPHVQKVVRELNRELGAQAPSLGDRTRLPYTDVAVLHE 359	
Db	299 FFAGTETVSTLRVYGFLLLMKYPEVEAKIHEINQVIGTHRTPRVDDRAKMPYTDVAVLHE 358	
QY	360 AORLLAVPMGIPTLMTTTRFRGYTLPGQTEVPFLGLSILHDPNIFKHPBEENPDRELD 419	
Db	359 IQRLTDIVELGVPHNVIRDTFRGYFLPKGTDVPLGSLVKDPKYFRYPEAFYQHFELD 418	
QY	420 ADGRFRKEAFELPFSLGRVCLGEGLAARLFLPFTTTLQAFSLSPCPDPTLSKPTVS 479	
Db	419 EQGRFKNDFAVAFSSGKRICVGEALARMELFLYFTSLQRFSLRSLVPPADIDIAHKIS 478	
QY	480 GLFNIPPAFOL 490	
Db	479 GFGNIPPTYEL 489	
RESULT 10		
CPB6 HUMAN		
ID	CPB6 HUMAN STANDARD; PRT; 491 AA.	
AC	P20813; Q9UK46;	
DT	01-FEB-1991 (Rel. 17, Created)	
DT	01-FEB-1991 (Rel. 17, Last sequence update)	
DT	10-OCT-2003 (Rel. 42, Last annotation update)	
DE	Cytochrome P450 2B6 [EC 1.14.14.1] (CYP2B6) (P450 IIB1).	
GN	CYP2B6.	
OS	Homo sapiens (Human).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
OX	NCBI_TaxID=9606;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	TISSUE=Liver;	
RX	MEDLINE=90057429; PubMed=2573390;	
RA	Yamano S., Nambu P.T., Aoyama T., Meyer U.A., Inaba T., Kalow W.,	
RA	Gelboin H.V., McBride O.W., Gonzalez F.J.;	
RT	"cDNA cloning and sequence and cDNA-directed expression of human P450	
RT	IIB1: identification of a normal and two variant cDNAs derived from	
RT	the CYP2B locus on chromosome 19 and differential expression of the	
RT	IIB mRNAs in human liver.";	
RL	Biochemistry 28:7340-7348(1989).	
RN	[2]	
RP	SEQUENCE FROM N.A., AND VARIANTS ALA-167; HIS-172 AND ARG-262.	
RC	TISSUE=Liver;	
RA	Zhuge J., Qian Y., Xie H., Yu Y.;	
RT	"Sequence of a new human cytochrome P450-2B6 cDNA.";	
RL	Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.	
RN	[3]	
RP	SEQUENCE FROM N.A.	
RA	Lamerdin J.E., McCready P.M., Skowronski E., Viswanathan V.,	
RA	Burkhardt-Schultz K., Gordon L., Dias J., Ramirez M., Stilwagen S.,	
RA	Phan H., Velasco N., Do L., Regala W., Terry A., Brower A., Ganes J.,	
RA	Danganan L., Exler A., Christensen M., Georgescu A., Avila J., Liu S.,	
RA	Andreise T., Frankheim M., Attix C., Amico-Keller G., Coefield J.,	
RA	Duarte S., Lucas S., Bruce R., Thomas P., Quan G., Krommiller B.,	
RA	Arellano A., Sanders C., Ow D., Nolan M., Trong S., Kobayashi A.,	
RA	Olsen A.S., Carrano A.V.;	
RT	"Sequence analysis of a 4-Mb region in 19q13.2.";	
RL	Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.	
RN	[4]	
RP	TISSUE SPECIFICITY.	
RX	MEDLINE=20228924; PubMed=10768437;	
RA	Thum T., Borlak J.;	
RT	"Gene expression in distinct regions of the heart.";	



OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_taxID=10090;  
 RN [1]  
 RN TISSUE=Kidney;  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Kidney;  
 RX MEDLINE=8917954; PubMed=2703500;  
 RA Lindberg R., Burkhardt B., Ichikawa T., Negishi M.;  
 RT "The structure and characterization of type I P-450(15) alpha gene as  
 RT major steroid 15 alpha-hydroxylase and its comparison with type II P-  
 RT 450(15) alpha gene.";  
 RL J. Biol. Chem. 264:6465-6471 (1989).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=17NC/Z;  
 RA Jounaidi Y.;  
 RT "CDNA and amino acid sequence of a new cyp2a isoform overexpressed in  
 RT chemically induced mouse hepatoma.";  
 RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP MUTAGENESIS.  
 RX MEDLINE=89281737; PubMed=2733794;  
 RA Lindberg R., Negishi M.;  
 RT "Alteration of mouse cytochrome P450cch substrate specificity by  
 RT mutation of a single amino-acid residue.";  
 RL Nature 339:632-634 (1989).  
 RN [4]  
 RP TISSUE SPECIFICITY.  
 RX MEDLINE=99421934; PubMed=1040589;  
 RA Lavery D.J., Lopez-Molina L., Margueron R., Fleury-Olela F.,  
 RA Conquet F., Schibler U., Bonfils C.;  
 RT "Circadian expression of the steroid 15 alpha-hydroxylase (Cyp2a4) and  
 RT coumarin 7-hydroxylase (Cyp2a5) genes in mouse liver is regulated by  
 RT the PAR leucine zipper transcription factor DBP.";  
 RL Mol. Cell. Biol. 19:6488-6499 (1999).  
 CC -!- FUNCTION: EXHIBITS A HIGH COUMARIN 7-HYDROXYLASE ACTIVITY.  
 CC -!- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +  
 CC oxidized flavoprotein + H(2)O.  
 CC -!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.  
 CC -!- TISSUE SPECIFICITY: LIVER, WITH A STRONG CIRCADIAN RHYTHMICITY.  
 CC -!- CIRCADIAN EXPRESSION IS REGULATED BY DBP.  
 CC -!- DEVELOPMENTAL STAGE: IN LIVER; ACTIVITY 6 FOLD HIGHER IN FEMALES  
 CC THAN IN MALES.  
 CC -!- MISCELLANEOUS: There are only 11 differences between the sequence  
 CC of testosterone 15-alpha-hydroxylase and that of coumarin 7-  
 CC hydroxylase. By site-directed mutagenesis it has been shown that  
 CC modification of position 209 is sufficient to convert the  
 CC specificity of the two forms of the enzyme.  
 CC -!- SIMILARITY: Belongs to the cytochrome P450 family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M25211; AAA37798.1; .  
 DR EMBL; M25204; AAA37798.1; JOINED.  
 DR EMBL; M25205; AAA37798.1; JOINED.  
 DR EMBL; M25206; AAA37798.1; JOINED.  
 DR EMBL; M25207; AAA37798.1; JOINED.  
 DR EMBL; M25208; AAA37798.1; JOINED.  
 DR EMBL; M25209; AAA37798.1; JOINED.  
 DR EMBL; M25210; AAA37798.1; JOINED.  
 DR EMBL; X89864; CAA61963.1; .  
 DR PIR; B33531; B33531.  
 DR HSSP; P00179; 1DT6.  
 DR MGD; MGI:88597; Cyp2a5.  
 DR InterPro; IPR001128; Cytochrome P450.  
 DR InterPro; IPR008067; EP450\_Cyp2A.

DR Pfam: PF00067; P450; 1.  
 DR PRINTS; PR01684; EP450ICYP2A.  
 DR PRINTS; PR00385; P450.  
 DR PROSITE; PS00086; CYTOCHROME\_P450; 1.  
 KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;  
 KW Microsome; Endoplasmic reticulum.  
 FT METAL 439 439 IRON (HEME AXIAL LIGAND).  
 SQ SEQUENCE 494 AA; 56740 MW; 1C2516D5FA2551D0 CRC64;  
 Query Match 47.3%; Score 1237; DB 1; Length 494;  
 Best Local Similarity 48.9%; Pred. No. 4.7e-79;  
 Matches 239; Conservative 94; Mismatches 150; Indels 6; Gaps 3;  
 QY 6 TWALLALALLALLLTLALSGT----RARGHLPPTPLPLLLGNLLQLRPGALYSGIMRLS 61  
 DB 3 TSGLLLVAAVFLSVLMSVWKQKLSGKLPPTPLPFFIGNLQNTQMYNSLMKIS 62  
 QY 62 KYQGVFTIYIGMPVVLVQGEAVREALGQQAEEFSGRGIVAMLEGTFDGHGVFSSG 121  
 DB 63 QRYGVFTIYILGP-RRIVVLCQGEAVKALVDQAEEFSGRGQATFDMLFKYGIVGVVFSG 121  
 QY 122 EHWROLRKFTMLALRDGLGKREGELIQAEARCLVETPQGTGPRFPDPSLLIAQATSNV 181  
 DB 122 ERAKQLRRPSTATLRDFGVGKKGIEERIQEEAGFLIDSPKTINGAFIDFTFYLSTVSNV 181  
 QY 182 VCSLLFLGLRFSYEDKEFOAVVRAAGTLLGVSSQGGQTYEMFSWFLRPLPGPHKQLLHV 241  
 DB 182 ISSIVFGDFDYEDKEFLSLRLMLGSGFQTATSMGLQYEMFSSVMKHLPGPQQQAFKEL 241  
 QY 242 STLAAPTQVQVQOQGNLDASGAPARDLVDAFLKKAQEQNPCTETNKNMLMTVLVLLF 301  
 DB 242 QGLEDFITKKVHNORTLDPNSP-RDFIDSLIRMLEKKNPENTEFYMKNLVLTUNLFF 300  
 QY 302 AGTMVSTTVGTLLLLMKYPHVQKWVRELARELGAGQAPSLGDRTRLPYTDVILHEAQ 361  
 DB 301 AGTETVSTTLRYGFLLLMKHPDIEAKVHEEIDRVIGRNQPKYEDRMKMPYEAIVHEIQ 360  
 QY 362 RLLALVPMGIPRTLMRTTRFRGYTLFPQGTVEFPLGSLIHLDPNIFKHPBEFNPDRFLDAD 421  
 DB 361 RADMLPMPGLARRVTKDTRFDFLLPKGTVEFPMGLSVLKDPKFPSPKDFNPKHFLDDK 420  
 QY 422 GPRKRHEAPLPSLGRVCLGEGALAKAELFLFTTTLQAFSLSPSPPTDLSLKPTVSG 481  
 DB 421 GQPKKNDAPVPSIGKRYCFEGELARMELFLFTNTIMQNFHFKSTQAPQDIDVSPRLVGF 480  
 QY 482 FNIPPAFQL 490  
 DB 481 ATIPPTITM 489  
 RESULT 12  
 CPBI\_MOUSE STANDARD; PRT; 492 AA.  
 ID CPBI\_MOUSE  
 AC OS5071;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Cytochrome P450 2B19 (EC 1.14.14.1) (CYP11B19).  
 GN CYP2B19.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6; TISSUE=Skin;  
 RX MEDLINE=99017986; PubMed=9799616;  
 RA Keeney D.S.;  
 RT "The novel skin-specific cytochrome P450 Cyp2b19 maps to proximal  
 RT chromosome 7 in the mouse, near a cluster of Cyp2 family genes.";  
 RL Genomics 53:417-419 (1998).  
 CC -!- FUNCTION: Cytochromes P450 are a group of heme-thiolate  
 CC monooxygenases. In liver microsomes, this enzyme is involved in an



CC	EMBL; M26208; AAA37797.1; .	DT	01-FEB-1991 (Rel. 17, last sequence update)
DR	EMBL; M25146; AAA37797.1; JOINED.	DT	28-FEB-2003 (Rel. 41, last annotation update)
DR	EMBL; M25147; AAA37797.1; JOINED.	DE	Cytochrome P450 2A3 (EC 1.14.14.1) (CYPIIA3) (Coumarin 7-hydroxylase).
DR	EMBL; M26202; AAA37797.1; JOINED.	GN	CYP2A3 OR CYP2A-3. (Rat).
DR	EMBL; M26203; AAA37797.1; JOINED.	OS	Rattus norvegicus (Rat).
DR	EMBL; M26205; AAA37797.1; JOINED.	OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
DR	EMBL; M26206; AAA37797.1; JOINED.	OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
DR	EMBL; M26207; AAA37797.1; JOINED.	OX	NCBI_taxID=10116;
DR	EMBL; J03549; AAA40426.1; .	RN	[1]
DR	EMBL; M19319; AAA40429.1; .	RP	SEQUENCE FROM N.A.
DR	PIR; A33531; A33531.	RC	STRAIN=Sprague-Dawley; TISSUE=Lung;
DR	HSSP; P00179; 1DT6.	RC	MEDLINE=90356430; PubMed=2388852;
DR	MGI; M8596; Cyp2a4.	RX	Ueno T., Gonzalez F.;
DR	InterPro; IPR001128; Cytochrome P450.	RA	"Complete sequence of the rat CYP2A3 gene specifically transcribed in
DR	InterPro; IPR008067; EP450_CYP2A.	RT	lung.";
DR	Pfam; PF00067; P450.1	RL	Nucleic Acids Res. 18:4623-4624(1990).
DR	PRINTS; PR01684; EP450ICYP2A.	RN	[2]
DR	PROSITE; PS00086; CYTOCHROME P450; 1.	RP	SEQUENCE OF 8-494 FROM N.A.
KW	Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;	RX	TISSUE=Lung;
KW	Microsome; Endoplasmic reticulum.	CC	MEDLINE=89323084; PubMed=2751996;
FT	METAL 439 439 IRON (HEME AXIAL LIGAND) .	RA	Kimura S., Kozak C.A., Gonzalez F.J.;
FT	CONFLICT 236 236 L -> Q (IN REF. 2).	RT	"Identification of a novel P450 expressed in rat lung: cDNA cloning
FT	CONFLICT 306 306 V -> G (IN REF. 2).	RT	and sequence, chromosome mapping, and induction by
SQ	SEQUENCE 494 AA; 56594 MW; 69724BDE3195D75C CRC64;	RT	3-methylcholanthrene.";
		RL	Biochemistry 28:3798-3803(1989).
		CC	-!- FUNCTION: Cytochromes P450 are a group of heme-thiolate
		CC	monooxygenases. In liver microsomes, this enzyme is involved in an
		CC	NADPH-dependent electron transport pathway. It oxidizes a variety
		CC	of structurally unrelated compounds, including steroids, fatty
		CC	acids, and xenobiotics.
		CC	-!- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +
		CC	oxidized flavoprotein + H(2)O.
		CC	-!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
		CC	-!- TISSUE SPECIFICITY: Lung.
		CC	-!- INDUCTION: By 3-methylcholanthrene (3MC).
		CC	-!- SIMILARITY: Belongs to the cytochrome P450 family.
		CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
		CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
		CC	the European Bioinformatics Institute. There are no restrictions on its
		CC	use by non-profit institutions as long as its content is in no way
		CC	modified and this statement is not removed. Usage by and for commercial
		CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/
		CC	or send an email to license@isb-sib.ch).
		CC	-----
		CC	EMBL; M33190; AAA41022.1; .
DR	EMBL; J02852; AAA88511.1; .	DR	EMBL; J02852; AAA88511.1; .
DR	HSSP; P00179; 1DT6.	DR	HSSP; P00179; 1DT6.
DR	InterPro; IPR001128; Cytochrome P450.	DR	InterPro; IPR001128; Cytochrome P450.
DR	InterPro; IPR008067; EP450_CYP2A.	DR	InterPro; IPR008067; EP450_CYP2A.
DR	Pfam; PF00067; P450.1.	DR	Pfam; PF00067; P450.1.
DR	PRINTS; PR01684; EP450ICYP2A.	DR	PRINTS; PR01684; EP450ICYP2A.
DR	PROSITE; PS00086; CYTOCHROME P450; 1.	DR	PROSITE; PS00086; CYTOCHROME P450; 1.
KW	Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;	KW	Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
KW	Microsome; Endoplasmic reticulum.	KW	Microsome; Endoplasmic reticulum.
FT	METAL 439 439 IRON (HEME AXIAL LIGAND) .	FT	METAL 439 439 IRON (HEME AXIAL LIGAND) .
FT	CONFLICT 109 109 W -> G (IN REF. 2).	FT	CONFLICT 109 109 W -> G (IN REF. 2).
FT	CONFLICT 335 335 I -> N (IN REF. 2).	FT	CONFLICT 335 335 I -> N (IN REF. 2).
SQ	SEQUENCE 494 AA; 56510 MW; F71855CBE602672F CRC64;	SQ	SEQUENCE 494 AA; 56510 MW; F71855CBE602672F CRC64;
		Query Match	47.2%; Score 1234.5; DB 1; Length 494;
		Best Local Similarity	49.1%; Pred. No. 6.4e-79;
		Matches 240; Conservative	91; Mismatches 152; Indels 6; Gaps 3;
QY	6 TWALLALALLLLLTALSGT-----RARGHLPPGPTPLPLGNLLQLRPGALYSGLMRLS 61	QY	6 TWALLALALLLLLTALSGT-----RARGHLPPGPTPLPLGNLLQLRPGALYSGLMRLS 61
DB	3 TSGLLVAFAVLSVLMSVWKQKLSGLKPLPGPTPLPLFVGNLQNTQMYNSLMKTS 62	DB	3 TSGLLVAFAVLSVLMSVWKQKLSGLKPLPGPTPLPLFVGNLQNTQMYNSLMKTS 62
QY	62 KYGVPFTIYLGPWPVVLVGOEAVREALGQAEFSGRGVAMLEGTFDGHGVFSSG 121	QY	62 KYGVPFTIYLGPWPVVLVGOEAVREALGQAEFSGRGVAMLEGTFDGHGVFSSG 121
DB	63 QYGGPFTIYLGS-RRVVLGQAEVKEALVQAEFSGRGQATFDWLFKGVGIAFSSG 121	DB	63 QYGGPFTIYLGS-RRVVLGQAEVKEALVQAEFSGRGQATFDWLFKGVGIAFSSG 121
QY	122 ERWRQLRKTMLALRDLMGKREGBELIQAEARCLVETTFQGTGRPPDFPSLLAQATSNV 181	QY	122 ERWRQLRKTMLALRDLMGKREGBELIQAEARCLVETTFQGTGRPPDFPSLLAQATSNV 181
DB	122 ERAKQLRSFSIATLDFGVGKRGIEBRIQEEAGFLDLSFRKNGAFIDPTFYLSTVSNV 181	DB	122 ERAKQLRSFSIATLDFGVGKRGIEBRIQEEAGFLDLSFRKNGAFIDPTFYLSTVSNV 181
QY	182 VCSLLPGLRFSYEDKQFAVRAAGTLLGVSSQGGQVEMFSWFLRPLPGPHKQLLHV 241	QY	182 VCSLLPGLRFSYEDKQFAVRAAGTLLGVSSQGGQVEMFSWFLRPLPGPHKQLLHV 241
DB	182 ISSIVGDRFDYEDKFEFLSLRRMLGSLQFTATSMGQVYEMFSSVMKHLPGQQQAFKEL 241	DB	182 ISSIVGDRFDYEDKFEFLSLRRMLGSLQFTATSMGQVYEMFSSVMKHLPGQQQAFKEL 241
QY	242 STLAFTVRQVOHOGNLDASGARDLVDAFLKMAQEQNPGTFTNKMMLTVIYLLF 301	QY	242 STLAFTVRQVOHOGNLDASGARDLVDAFLKMAQEQNPGTFTNKMMLTVIYLLF 301
DB	242 QGLEDFITKVEHNTLDNSP-RDFIDSFLRMLEEKKNPTEFYMKNVLVLTTLNLF 300	DB	242 QGLEDFITKVEHNTLDNSP-RDFIDSFLRMLEEKKNPTEFYMKNVLVLTTLNLF 300
QY	302 AGTMTVSTTGYTLLLMKYPHVQKVRRELNGAGQAPSLGDRTRLPDYDAVLHEAQ 361	QY	302 AGTMTVSTTGYTLLLMKYPHVQKVRRELNGAGQAPSLGDRTRLPDYDAVLHEAQ 361
DB	301 AGTETVSTILRYGFLLLMKYPDIEAKVHEBIDRVIGNRQPKYEDRMKMPYTEAVIHEIQ 360	DB	301 AGTETVSTILRYGFLLLMKYPDIEAKVHEBIDRVIGNRQPKYEDRMKMPYTEAVIHEIQ 360
QY	362 RLALVPMGIPRLMTRTRGTYLPGTVEFLLGSLILDHDPNIKHPESFNPDRFLDAD 421	QY	362 RLALVPMGIPRLMTRTRGTYLPGTVEFLLGSLILDHDPNIKHPESFNPDRFLDAD 421
DB	361 RFADLIPMGILARRVTKDTRFDLPLPKGTVEFVPMGLSVLKDPPKDFNPKHFLDDK 420	DB	361 RFADLIPMGILARRVTKDTRFDLPLPKGTVEFVPMGLSVLKDPPKDFNPKHFLDDK 420
QY	422 GRFRKEAFPLSFGKRVCLGEGLAKEALFFFTTLTQAFSLSPSPCPDPTLSLKPTVSGL 481	QY	422 GRFRKEAFPLSFGKRVCLGEGLAKEALFFFTTLTQAFSLSPSPCPDPTLSLKPTVSGL 481
DB	421 GQPKKSDAFVFFSFGKRVCFGEGLARMELFLFLTNMQNHFKSTQAPQIDVSPRLVGF 480	DB	421 GQPKKSDAFVFFSFGKRVCFGEGLARMELFLFLTNMQNHFKSTQAPQIDVSPRLVGF 480
QY	482 FNIPPAFQL 490	QY	482 FNIPPAFQL 490
DB	481 VTIPPTYTM 489	DB	481 VTIPPTYTM 489
		RESULT 14	
		CPA3_RAT	
		ID_CPA3_RAT	
		AC_P20812;	
		DT 01-FEB-1991 (Rel. 17, Created)	



[illegible]

RESULT 15

CPAD_HUMAN	
ID	CPAD_HUMAN
AC	Q6696; Q9H2X2;
QC	15-DEC-1998 (Rel. 37, Created)
DT	15-DEC-1998 (Rel. 37, Last sequence update)
DT	10-OCT-2003 (Rel. 42, Last annotation update)
DE	Cytochrome P450 2A13 (EC 1.14.14.1) (CYP11A13).
GN	CYP2A13.
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID	9606;
LN	[1]
RP	SEQUENCE FROM N.A.
RX	MEDLINE=95337851; PubMed=7668294;
RA	Fernandez-Salguero P., Hoffman S.M., Cholerton S., Mohrenweiser H.,
RA	Raunio H., Rautio A., Pelkonen O., Huang J.D., Evans W.E., Idle J.R.,
RT	"A genetic polymorphism in coumarin 7-hydroxylation: sequence of the
RT	human CYP2A genes and identification of variant CYP2A6 alleles.";
RT	Am. J. Hum. Genet. 57:651-660(1995).
LN	[2]
RP	SEQUENCE FROM N.A.; CHARACTERIZATION, AND TISSUE SPECIFICITY.
RX	MEDLINE=20468846; PubMed=11016631;
RA	Su T., Bao Z., Zhang Q.Y., Smith T.J., Hong J.Y., Ding X.;
RT	"Human cytochrome P450 CYP2A13: predominant expression in the
RT	respiratory tract and its high efficiency metabolic activation of a
RT	tobacco-specific carcinogen, 4-(methylnitrosamino)-1-(3-pyridyl)-1-
RT	butanone.";
RT	Cancer Res. 60:5074-5079(2000).
LN	[3]
RP	VARIANT CYS-257.
RX	MEDLINE=22126112; PubMed=12130698;
RA	Zhang X., Su T., Zhang Q.Y., Gu J., Caggana M., Li H., Ding X.;
RT	"Genetic polymorphisms of the human CYP2A13 gene: identification of
RT	single-nucleotide polymorphisms and functional characterization of an
RT	Arg257Cys variant.";
RT	J. Pharmacol. Exp. Ther. 302:416-423(2002).
CC	-!- FUNCTION: Exhibits a coumarin 7-hydroxylase activity. Active in
CC	the metabolic activation of hexamethylphosphoramide, N,N-
CC	dimethylaniline, 2'-methoxyacetophenone, N-
CC	nitrosomethylphenylamine, and the tobacco-specific carcinogen, 4-
CC	(methylnitrosamino)-1-(3-pyridyl)-1-butanone.

us-10-669-693-2.rsp

Mon Sep 20 09:09:28 2004

Qy	357	LHEAQRLLALVPMGIPRTLMRTTRFRGYTLPQGTVEFPLGSLILHDPNIFKHPDEEFPDR	416
Db	356	IHEIQRTGDMPLMGLAHRVNDTKFRDFFLPKGTVEFPLGSELDRPRFFSNPDGSPQH	415
Qy	417	FLDADGRKHEAFLPESLGRVCLGEGAKAELFLPFTTILQAFSLSPCPDPTLSLKP	476
Db	416	FLDEKGQFKSDAFVPSIGKRYCFEGELARWELFLFTTIMQNERFKSPQSPKXIDVSP	475
Qy	477	TVSGLFNIPPAPQLQVRP	494
Db	476	KHVGFAIPRNYTMSFLP	493

Search completed: September 15, 2004, 09:19:00  
Job time : 26 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2004, 09:14:41 ; Search time 43 seconds  
(without alignments)  
1127.454 Million cell updates/sec

Title: US-10-669-693-2

Perfect score: 2615

Sequence: 1 MEATGTWALLALALLLLT.....PPAFOLQVRPTDLHSTTQTR 504

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR 78.\*

2: PIR2.\*

3: PIR3.\*

4: PIR4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2615	100.0	504	2 JC7613	cytochrome P450 2S
2	1319.5	50.5	494	2 B31944	cytochrome P450 2G
3	1313	50.2	491	2 S35666	cytochrome P450 2B
4	1300	49.7	491	1 O4RBP4	cytochrome P450 2B
5	1299	49.7	491	2 S31277	cytochrome P450 2B
6	1298	49.6	491	2 S31277	cytochrome P450 2B
7	1296	49.6	491	2 S31278	cytochrome P450 2B
8	1293	49.4	491	1 O4RTP2	cytochrome P450 2B
9	1289	49.3	491	1 O4RTPB	cytochrome P450 2B
10	1280.5	49.0	500	2 B31047	cytochrome P450 2B
11	1275	48.8	491	2 JT0676	testosterone 16alp
12	1262	48.3	494	2 S11305	cytochrome P450 2B
13	1247	47.7	491	2 I84735	cytochrome P450 2B
14	1244	47.6	491	2 A31047	testosterone 16a-h
15	1243.5	47.6	494	2 A35551	testosterone 16alp
16	1238	47.3	491	2 A32969	cytochrome P450 2G
17	1237	47.3	494	2 B33531	cytochrome P450 2B
18	1235	47.2	494	2 A32030	testosterone 15alp
19	1234.5	47.2	494	2 A32030	cytochrome P450 2A
20	1220	46.7	494	2 S16068	testosterone 15alp
21	1213.5	46.4	494	1 O4HUA6	testosterone 16a-h
22	1210.5	46.3	492	2 S27160	cytochrome P450 2A
23	1210	46.3	491	2 I49625	cytochrome P450 2A
24	1203.5	46.0	494	2 A47494	cytochrome P450 2A
25	1198.5	45.8	494	2 B47494	cytochrome P450 -
26	1175.5	45.0	494	2 I38965	cytochrome P450 2A
27	1174.5	44.9	494	2 C34271	cytochrome P450 50
28	1168	44.7	491	2 A39302	cytochrome P450 ar
29	1163.5	44.5	494	2 A46588	

30 1160 44.4 491 2 A28818 cytochrome P450 2B  
31 1153.5 44.1 494 2 I38967 cytochrome P450 -  
32 1150.5 44.0 492 2 S32491 testosterone 7alpb  
33 1144.5 43.8 492 2 A34272 testosterone 7alpb  
34 1108 42.4 491 2 A36036 cytochrome P450 2F  
35 1107 42.3 491 2 A31418 cytochrome P450 2H  
36 1105.5 42.3 494 2 A33293 cytochrome P450 2A  
37 1098 42.0 487 1 OARBC6 cytochrome P450 2C  
38 1097.5 42.0 490 2 B28516 progesterone mono  
39 1096 41.9 487 1 OARBP4 testosterone 7alpb  
40 1091.5 41.7 492 2 A31887 testosterone 7alpb  
41 1090.5 41.7 494 2 S13101 progesterone mono  
42 1085.5 41.5 490 2 B38462 testosterone 7alpb  
43 1083 41.4 491 2 A34814 S-mephenytoin 4-hy  
44 1079.5 41.3 490 2 I48163 cytochrome P450 CY  
45 1078.5 41.2 490 2 D28951 cytochrome P450 -  
46 1078.5 41.2 490 2 D28951 cytochrome P450 2C

## ALIGNMENTS

### RESULT 1

JC7613

cytochrome P450 2S1 protein, CYP2S1 - human

C:Species: Homo sapiens (man)

C:Date: 30-Jun-2001 #sequence\_revision 30-Jun-2001 #text\_change 03-Jun-2002

C:Accession: JC7613

R:Rylander, T.; Neve, E.P.A.; Ingelman-Sundberg, M.; Oscarson, M.

Biochem. Biophys. Res. Commun. 281, 529-535, 2001

A:Title: Identification and tissue distribution of the novel human cytochrome P450 2S1 ((

A:Reference number: JC7613; MUID:21092856; PMID:11181079

A:Contents: Liver

A:Accession: JC7613

A:Molecule type: mRNA

A:Residues: 1-504 <RYL>

A:Cross-references: GB:AF335278

C:Comment: This protein, a novel member of cytochrome P450 2(CYP2) family, has a role in

C:Genetics:

A:Gene: cyp2s1

A:Map position: 19q13.2

A:Introns: 59/3; 115/1; 165/1; 218/3; 278/3; 326/1; 388/3; 436/1

C:Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology

C:Keywords: heme; iron; liver; metalloprotein

F:33-36/Region: Conserved PPGP sequence #status predicted

F:440/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 100.0%; Score 2615; DB 2; Length 504;  
Best Local Similarity 100.0%; Pred. No. 3.5e-205;  
Matches 504; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MEATGTWALLALALLLLTALS	GTARGHLPPGPTPLPLGNLLQLRPGALYSGLMRL	60
DB	1	MEATGTWALLALALLLLTALS	GTARGHLPPGPTPLPLGNLLQLRPGALYSGLMRL	60
QY	61	SKKYGVFTIYLGPRWPVVVL	VGQEAVRALGQAQAEFSGRGTVAMLEGTFDGHGVFFSN	120
DB	61	SKKYGVFTIYLGPRWPVVVL	VGQEAVRALGQAQAEFSGRGTVAMLEGTFDGHGVFFSN	120
QY	121	GERWRQLRFTMLALRDL	CGMKREGELIOAARCLVETFOGTGPRFPDPSLLAQTNS	180
DB	121	GERWRQLRFTMLALRDL	CGMKREGELIOAARCLVETFOGTGPRFPDPSLLAQTNS	180
QY	181	VVCSLLFGLRFSYEDKE	FQAVVRAAGTLLGVSSGGQTYEMFNSFLRPLPGPKQLLHH	240
DB	181	VVCSLLFGLRFSYEDKE	FQAVVRAAGTLLGVSSGGQTYEMFNSFLRPLPGPKQLLHH	240
QY	241	VSTLAFTVQVQOQHQGNL	DASGPARDLVDAFLKMAQEQNPQTEFTNKMLMTVIYLL	300
DB	241	VSTLAFTVQVQOQHQGNL	DASGPARDLVDAFLKMAQEQNPQTEFTNKMLMTVIYLL	300
QY	301	FAGTMTVSTTGYTTL	LLLMKYPHVQKVRRELNRRLGAGQAPSLGDRTRPLPYTDAVLHEA	360
DB	301	FAGTMTVSTTGYTTL	LLLMKYPHVQKVRRELNRRLGAGQAPSLGDRTRPLPYTDAVLHEA	360

Db 359 IQRLTIDVPMGVPHWIRDTFRGYLLPKGTVDVPLLGSLVKDKPKYFCHPDPDFQHQFLD 418  
 Qy 420 ADGRFKHEAFPLPSLGRKVCLEGLAKAEFLFETTLIOAFSLSPCCPDDTLSLKPTVS 479  
 Db 419 EQGRFKNEAFVPPSSKRICLEGEAMARMELFLFTSLQNFSLHPLVFPVNDITPKIS 478  
 Qy 480 GLFNIPPAFOL 490  
 Db 479 GFGNIPPTIEL 489

RESULT 3  
 S35666  
 Cytochrome P450 2B4 isoform Bx - rabbit  
 N/Alternate names: cytochrome P450 2B isoform Bx  
 N/Contains: oxidoreductase (EC 1.-.-.-)  
 C/Species: Oryctolagus cuniculus (domestic rabbit)  
 C/Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 04-Mar-2000  
 C/Accession: S35666; B54251  
 R/Ryan, R.; Grimm, S.W.; Kedzie, K.M.; Halpert, J.R.; Philpot, R.M.  
 Arch. Biochem. Biophys. 304, 454-463, 1993  
 A/Title: Cloning, sequencing, and functional studies of phenobarbital-inducible forms of  
 A/Reference number: S35666; MUID:93349061; PMID:8346920  
 A/Accession: S35666  
 A/Molecule type: mRNA  
 A/Residues: 1-491 <RYA>  
 A/Cross-references: EMBL:S64259; NID:G402842; PIDN:AAB27705.1; PID:G402843  
 A/Experimental source: kidney  
 R/Roberts, E.S.; Hopkins, N.E.; Zaluzec, E.J.; Gage, D.A.; Alworth, W.L.; Hollenberg, P.  
 Biochemistry 33, 3766-3771, 1994  
 A/Title: Identification of active-site peptides from (3)H-labeled 2-ethynylaphthalene-1  
 A/Reference number: A54251; MUID:94190699; PMID:8142377  
 A/Accession: B54251  
 A/Status: preliminary  
 A/Molecule type: protein  
 A/Residues: 273-283 <ROB>  
 C/Genetics:  
 A/Gene: CYP2B4  
 C/Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology  
 C/Keywords: chromoprotein; electron transfer; endoplasmic reticulum; heme; iron; metallo  
 F/295-458/Domain: cytochrome P450 homology <P45>  
 F/436/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 50.2%; Score 1313; DB 2; Length 491;  
 Best Local Similarity 50.7%; Pred. No. 4.6e-99;  
 Matches 247; Conservative 99; Mismatches 137; Indels 4; Gaps 3;

Qy 7 WALLALALL--LLLTALSGTRAGHLPPGTPPLGLNLLQRLPGALYSGLMLRSKY 64  
 Db 3 FSELLLLAFLAGLLLLFRGHKPAHGRLLPPGPSPLVGLNLLQMDRKLRLRSFLRREKY 62  
 Qy 65 GPVFTYILGPWRPVVVLVQEAAREALGQAEFEFGRTVAMLEGTFDGHGVFSGNGRW 124  
 Db 63 GDVFTVYLSG-RPVVVLGGTDAAREALVQEAAREALGQAEFEFGRTVAMLEGTFDGHGVFSGNGRW 121  
 Qy 125 RQLRKFTMLALRDLGMRKREBELIQAERCLVETFTQTEGRFPDPSLLAQAQTSNVVCS 184  
 Db 122 RALRRFSLATRDPFGMRKRSVEERIQAERCLVETFTQTEGRFPDPSLLAQAQTSNVVCS 181  
 Qy 185 LLEGLRFSYEDKFEQAVVRAAGTLLGVSSQGGQTYEMFSLFPLPGPHKQLLHVSTL 244  
 Db 182 IVEGKRDYKDPVFLRLDLLFQSFSLISSFSQVFEFLGFKHFFGTHQIRYRNIQEI 241  
 Qy 245 AAFTRVOOQHOGLNDASGPARDLVDAFLKMAEQEONPGTEFTKNMLMTVIYLLFAGT 304  
 Db 242 NTFITQSVKHEATLDPSPN-RDFIDVYLRLMEKDKSDPSSEFHQNLILTVLSFPAGT 300  
 Qy 305 MTVSTTVGYTLLLMKYPHVQKVRRELNRRELGAQPSLGDRTLRPLPYTDAVLHEAQLL 364  
 Db 301 ETTSTTLRYGLMLKYPHYVTERVQKIEQVIGSHRRPPALDDRAKMPYTDVAVIHEIQLG 360  
 Qy 365 ALVPMGIPRLMRTTRFRGYTLPGQTEVFFLLGSLILHDPNIFKHPPEEFPDRFLD 424

Qy 361 QRLALVPMGIPRLMRTTRFRGYTLPGQTEVFFLLGSLILHDPNIFKHPPEEFPDRFLD 420  
 Db 361 QRLALVPMGIPRLMRTTRFRGYTLPGQTEVFFLLGSLILHDPNIFKHPPEEFPDRFLD 420  
 Qy 421 DGRFKHEAFPLPSLGRKVCLEGLAKAEFLFETTLIOAFSLSPCCPDDTLSLKPTVS 480  
 Db 421 DGRFKHEAFPLPSLGRKVCLEGLAKAEFLFETTLIOAFSLSPCCPDDTLSLKPTVS 480  
 Qy 481 LFNIPPAFOLVPTDLHSTTQTR 504  
 Db 481 LFNIPPAFOLVPTDLHSTTQTR 504

RESULT 2  
 B31944  
 Cytochrome P450 2G1 - rabbit  
 N/Alternate names: cytochrome P450NMB  
 N/Contains: oxidoreductase (EC 1.-.-.-)  
 C/Species: Oryctolagus cuniculus (domestic rabbit)  
 C/Date: 31-Mar-1990 #sequence\_revision 17-May-1996 #text\_change 04-Mar-2000  
 C/Accession: S31907; B31944  
 R/Ding, X.; Porter, T.D.; Peng, H.M.; Coon, M.J.  
 Arch. Biochem. Biophys. 285, 120-125, 1991  
 A/Title: cDNA and derived amino acid sequence of rabbit nasal cytochrome P450NMB (P450I1  
 A/Reference number: S31907; MUID:91119408; PMID:1703755  
 A/Accession: S31907  
 A/Status: not compared with conceptual translation  
 A/Molecule type: mRNA  
 A/Residues: 12-494 <DIN>  
 R/Ding, X.; Coon, M.J.  
 Biochemistry 27, 8330-8337, 1988  
 A/Title: Purification and characterization of two unique forms of cytochrome P-450 from  
 A/Reference number: A90541; MUID:89207441; PMID:3242590  
 A/Accession: B31944  
 A/Molecule type: protein  
 A/Residues: 1-45, 'E', 47-50 <D12>  
 C/Genetics:  
 A/Gene: CYP2G1  
 C/Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology  
 C/Keywords: chromoprotein; electron transfer; endoplasmic reticulum; heme; iron; metallo  
 F/298-461/Domain: cytochrome P450 homology <P45>  
 F/439/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 50.5%; Score 1319.5; DB 2; Length 494;  
 Best Local Similarity 50.5%; Pred. No. 1.4e-99;  
 Matches 248; Conservative 97; Mismatches 143; Indels 3; Gaps 3;

Qy 1 MEATGWALLALAL--LLLTALSGTRAGHLPPGTPPLGLNLLQRLPGALYSGLMR 59  
 Db 1 MEIGGAFTIPLALCFSCLLILIAWKVKQKPRPPGTPPIFFGLNLLQVETDATTQSFLLK 60  
 Qy 60 LSKKYGPVFTYILGPWRPVVVLVQEAAREALGQAEFEFGRTVAMLEGTFDGHGVFSS 119  
 Db 61 LREKYGPVFTYILGPWRPVVVLVQEAAREALGQAEFEFGRTVAMLEGTFDGHGVFSS 119  
 Qy 120 NGRWRQLRFTMLALRDLGMRKREBELIQAERCLVETFTQTEGRFPDPSLLAQAQTS 179  
 Db 120 NGEWRLLRPSLTLLADFGMKSEIERIQEAGYLLFEEFRKTKGAPIDPTFFLSRTVS 179  
 Qy 180 NVVCSLLFGRLFSYEDKFEQAVVRAAGTLLGVSSQGGQTYEMFSLFPLPGPHKQLLHV 239  
 Db 180 NVISSVVGSRFDYEDKQFLSLRLMINESFIEMSTPWAQLYDMYGVNQVQLPGRHRIY 239  
 Qy 240 HVSTLAFTVROOQHOGLNDASGPARDLVDAFLKMAEQEONPGTEFTKNMLMTVIYLL 299  
 Db 240 LIEELKDFIARVKNVNEASLDPPN-RDFIDCFILKHKQDKNNPHTFENLKNLVLTLNL 298  
 Qy 300 LFAGTMTVSTTVGYTLLLMKYPHVQKVRRELNRRELGAQPSLGDRTLRPLPYTDAVLHE 359  
 Db 299 PFAGTETVSSTLRYGLMLKYPHYVTERVQKIEQVIGSHRRPPALDDRAKMPYTDVAVI 358  
 Qy 360 AQRLLALVPMGIPRLMRTTRFRGYTLPGQTEVFFLLGSLILHDPNIFKHPPEEFPDRFLD 419

```

Db 361 DIPIFGVHTVTKDQFRGVIPKNTVEPVVSSALHDPYFKTPNTFNGHFLDANGAL 420
QY 425 RKHEAFPLPSLCKRYCLGSLAKAELEFETTLQAFSLSPDPTLSLKTPTVSLFNI 484
Db 421 KNEGFMPFSLGKRYCLGSLGIAETELFETTLQAFSLSPDPTLSLKTPTVSLFNI 480
QY 485 PPAFQLQ 491
Db 481 PPSYQIR 487

RESULT 4
cytochrome P450 2B4 - rabbit
N:Alternate names: cytochrome P450-1M2
N:Contains: oxidoreductase (EC 1.-.-.-)
C:Species: Oryctolagus cuniculus (domestic rabbit)
C>Date: 20-Sep-1984 #sequence revision 20-Sep-1984 #text change 03-Mar-2000
C:Accession: A00179; A61538; S31279; A00178; B27717; C27717; E27717
R:Tarr, G.E.; Black, S.D.; Fujita, V.S.; Coon, M.J.
Proc. Natl. Acad. Sci. U.S.A. 80, 6552-6556, 1983
A:Title: Complete amino acid sequence and predicted membrane topology of phenobarbital-1
A:Reference number: A00179; MUID:84042509; PMID:6579541
A:Accession: A00179
A:Molecule type: protein
A:Residues: 1-491 <TAR>
R:Parandooch, Z.; Fujita, V.S.; Coon, M.J.; Philpot, R.M.
Drug Metab. Dispos. 15, 59-67, 1987
A:Title: Cytochrome P-450 isozymes 2 and 5 in rabbit lung and liver. Comparisons of stru
A:Reference number: A61538; MUID:87161284; PMID:2881760
A:Accession: A61538
A:Molecule type: protein
A:Residues: 1-24 <PAR>
R:Gasser, R.; Negishi, M.; Philpot, R.M.
Mol. Pharmacol. 33[32], 22-30, 1988
A:Title: Primary structures of multiple forms of cytochrome P-450 isozyme 2 derived from
A:Reference number: S31277
A:Note: header on page 22 gives volume number as 32
A:Accession: S31279
A:Molecule type: mRNA
A:Residues: 1-491 <GAS>
A:CROSS-references: EMBL:M20856; NID:g164958; PIDN:AAA65840.1; PID:g164959
R:Heinemann, F.S.; Ozols, J.
J. Biol. Chem. 258, 4195-4201, 1983
A:Title: The complete amino acid sequence of rabbit phenobarbital-induced liver microsom
A:Reference number: A00178; MUID:83160983; PMID:6833251
A:Accession: A00178
A:Molecule type: protein
A:Residues: 1-90, 'E', 92-94, 96, 95, 97-98, 101-134, 'GY', 137-140, 'G', 142-192, 'K', 194-220, 'S',
R:Komori, M.; Imai, Y.; Tsunashawa, S.; Sato, R.
Biochemistry 27, 73-80, 1988
A:Title: Microheterogeneity in the major phenobarbital-inducible forms of rabbit liver m
A:Reference number: A90538; MUID:88163620; PMID:2831964
A:Accession: B27717
A:Molecule type: mRNA
A:Residues: 301-313, 'L', 315-419, 'M', 421-491 <KOM>
A:Experimental source: clone b14
A:Accession: C27717
A:Molecule type: mRNA
A:Residues: 52-103, 'M', 105-173, 'V', 175-220, 'S', 222-313, 'L', 315-424, 'C', 426-491 <KO2>
A:Experimental source: clone b46
A:Note: the authors translated the codon AAG for residue 191 as Arg
A:Accession: E27717
A:Molecule type: mRNA
A:Residues: 297-479, 'L', 481-491 <KO3>
A:Experimental source: clone b54
C:Comment: Cytochromes P450 are a group of membrane-bound hemoprotein monooxygenases. In
variety of structurally unrelated compounds, including steroids, fatty acids, and xenob
C:Genetics:
A:Gene: CYP2B4
C:Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology
C:Keywords: chromoprotein; electron transfer; endoplasmic reticulum; heme; iron; metallo

```

F:295-458/Domain: cytochrome P450 homology <P45>  
F:436/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 49.7%; Score 1300; DB 1; Length 491;  
Best Local Similarity 50.1%; Pred. No. 5.3e-98;  
Matches 244; Conservative 101; Mismatches 138; Indels 4; Gaps 3;

```

QY 7 WALLALALL--LLLTALSGTRARGHLPFGTPTPLGLNLLQLRPGALYSGLMRLSKY 64
Db 3 FSLLLLAFLAGLLLLFRGHPKAHGRLLPFGPSPLVGLNLLQMDRKGLLRFLRLREKY 62
QY 65 GPVFTYILGPRPVVYVNGQEAVERALGQAEFSGRGTVAMLEGTFGHGVFFNSGERM 124
Db 63 GDFVTYILGS-RPVVVLGCTDAIREALVDQAEAFSGRGKIAVVDPIFGYGVIFANGERM 121
QY 125 RQLRFTMLALRDLGMRGEGEELIQAEARCLIVETFGTEGPRPDPSPSLLAQATSNVVC 184
Db 122 KALRFSLATWRDFGMRKRSVEERIQEAEARCLIVELRKSKGALLDNTLLFHSITSNICS 181
QY 185 LLFGLRFSYEDKEFOAVVRAAGTLLGVSSQGGQTYEMFSWFLRPLPGPHKOLLHHVSTL 244
Db 182 IVFGKRFYKDPVFLRLDLDFFQSFSLISSFSQVVELFPGFLKHFPGTHRQIYRNLOEI 241
QY 245 AAFTRVQVQHQGNLDASGPARDLVDAFLKMAQEEQNPTEFTNNMLMTVIYLLFAGT 304
Db 242 NTFIGQSVEKRRATLDPSNP-RDFIDVYLLRMEKDKSPSEFHHQNLILTVLSFFAGT 300
QY 305 MTVSTTVGYTLLMLKYPHVOKWVREELNRELGAQAPSLGDRTLPTVDAVLHQAQRL 364
Db 301 ETTTILRYGLMLKYPHVTERVQKIEQVIGSHRPPALDDRAKMPYTDVVIHIOIRLG 360
QY 365 ALVPMGIPRLMTRFRGVTLPQGTVEVFPVLLGSLILHDPNIFKHPEEFNPRDLADQRE 424
Db 361 DLIPFGVHTVTKDQFRGVIPKNTVEPVVSSALHDPYFKTPNTFNGHFLDANGAL 420
QY 425 RKHEAFPLPSLCKRYCLGSLAKAELEFETTLQAFSLSPDPTLSLKTPTVSLFNI 484
Db 421 KNEGFMPFSLGKRYCLGSLGIAETELFETTLQAFSLSPDPTLSLKTPTVSLFNI 480
QY 485 PPAFQLQ 491
Db 481 PPSYQIR 487

```

## RESULT 5

S31277  
cytochrome P450 2B4-B1 - rabbit  
N:Contains: oxidoreductase (EC 1.-.-.-)  
C:Species: Oryctolagus cuniculus (domestic rabbit)  
C>Date: 28-May-1993 #sequence revision 28-May-1993 #text change 04-Mar-2000  
C:Accession: S31277; I46699  
R:Gasser, R.; Negishi, M.; Philpot, R.M.  
Mol. Pharmacol. 33[32], 22-30, 1988  
A:Title: Primary structures of multiple forms of cytochrome P-450 isozyme 2 derived from  
A:Reference number: S31277  
A:Accession: S31277  
A:Molecule type: mRNA  
A:Residues: 1-491 <GAS>  
A:CROSS-references: EMBL:M20857; NID:g164960; PIDN:AAA31224.1; PID:g164961  
A:Gene: CYP2B4  
C:Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology  
C:Keywords: chromoprotein; electron transfer; endoplasmic reticulum; heme; iron; metallo  
F:295-458/Domain: cytochrome P450 homology <P45>  
F:436/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 49.7%; Score 1299; DB 2; Length 491;  
Best Local Similarity 50.1%; Pred. No. 6.4e-98;  
Matches 244; Conservative 101; Mismatches 138; Indels 4; Gaps 3;

```

QY 7 WALLALALL--LLLTALSGTRARGHLPFGTPTPLGLNLLQLRPGALYSGLMRLSKY 64
Db 3 FSLLLLAFLAGLLLLFRGHPKAHGRLLPFGPSPLVGLNLLQMDRKGLLRFLRLREKY 62

```





Db 63 GDFVTVLGS-RPVVVLGCTDAIREALVQDAEAFSGRGIADVDFIQGYGVFFANGEM 121  
 QY 125 ROLKFTMLARDLGMKREBELLOAERCLVETFGTEGPPDPSPSLLLAQATSNVCS 184  
 Db 122 RALRRFSUATWRDFGMGRKSVVEERIOEERCLVSELKSGALLDNTLLFHSVTNICS 181  
 QY 185 LLFGLRFSYEDKEQAVVRAAGGTLLGVSSGGQTYEMFSWFLRLPLPGPHKOLLHHVSTL 244  
 Db 182 IVFGKRPYKDPVFLRLDLDFQSFSLISSSSQVFLPFGFLKHPGTHQIYRNLOEI 241  
 QY 245 AAFVTRVQOQOQGNLDASGPARDLVDAFLKVAQEQNGTEFTNKNMLMTVILLFAGT 304  
 Db 242 NTFTGQTVKEHRATLDSPN-RDFIDVLLRMKXKSDPSSEFHRNLLITVLTLFFAGT 300  
 QY 305 MIVSTTVGYTLLMLKYPHVQKWRLEENRELKAGQAPSLGDRTRLPTDVAVLHHAQRLL 364  
 Db 301 ETTSTTIRYGFLLMLKYPHVQKWRLEENRELKAGQAPSLGDRTRLPTDVAVLHHAQRLL 360  
 QY 365 ALVPMGIPTLMTRTRFGTYLTPQGTVEVFLPGLSLHDPNIFKHPENPDRELDADGRF 424  
 Db 361 DLVPFGAPHMVTKDTQFRGVYIPKNTVEVFLPGLSLHDPNIFKHPENPDRELDADGRF 420  
 QY 425 RXHEAPLPSLGRKVCLEGLAKAELFFFTTILQAFSLSPCPDPTLSLKPVTSGLFNI 484  
 Db 421 KRNEGFMFSLGRKVCLEGLAKAELFFFTTILQAFSLSPCPDPTLSLKPVTSGLFNI 480  
 QY 485 PPAFQLQ 491  
 Db 481 PPSYQIR 487

RESULT 8  
 O4RTP2  
 Cytochrome P450 2B2 - rat  
 N;Alternate names: cytochrome P450 PB-4; cytochrome P450, phenobarbital-inducible; cytochrome P450 2B2  
 C;Species: Rattus norvegicus (Norway rat)  
 C;Date: 04-Dec-1986 #sequence revision 17-May-1996 #text change 01-Dec-2000  
 C;Accession: A21162; A00177; B00176; B92255; S15589; A21872; A32736; S03855; I59060  
 R;Mizukami, Y.; Sogawa, K.; Suwa, M.; Muramatsu, M.; Fujii-Kuriyama, Y.  
 Proc. Natl. Acad. Sci. U.S.A. 80, 3958-3962, 1983  
 A;Title: The structure of phenobarbital-inducible cytochrome P-450 in rat liver.  
 A;Reference number: A21162; MUID:83247397; PMID:6306654  
 A;Accession: A21162  
 A;Molecule type: DNA  
 A;Residues: 1-472, 'M', 474-491 <MTZ>  
 A;Cross-references: EMBL:J00728; NID:G203845; PIDN:AAA41056.1; PID:G203847  
 A;Note: the authors translated the codon AGT for residue 4 as Thr, and ATG for residue 3 as Thr.  
 R;Freij, A.B.; Waxman, D.J.; Kreibich, G.  
 J. Biol. Chem. 260, 15253-15265, 1985  
 A;Title: The structure of phenobarbital-inducible rat liver cytochrome P-450 isoenzyme P-450  
 A;Reference number: A00177; MUID:86059379; PMID:387725  
 A;Accession: A00177  
 A;Molecule type: protein  
 A;Residues: 1-291, 'P', 293-320, 'AE', 323-475, 'D', 477-491 <PRE>  
 R;Fujii-Kuriyama, Y.; Mizukami, Y.; Kawajiri, K.; Sogawa, K.; Muramatsu, M.  
 Proc. Natl. Acad. Sci. U.S.A. 79, 2793-2797, 1982  
 A;Title: Primary structure of a cytochrome p-450: coding nucleotide sequence of phenobarbital  
 A;Reference number: A93912; MUID:82222224; PMID:6953431  
 A;Accession: B00176  
 A;Molecule type: mRNA  
 A;Residues: 6-359, 'S', 361-362, 'V', 364-366, 'V', 368-406, 'S', 408-416, 'N', 418, 'A', 420-477, 'G'  
 A;Note: nucleotide sequence for residues 1-5 is not given  
 R;Fujii-Kuriyama, Y.; Mizukami, Y.; Kawajiri, K.; Sogawa, K.; Muramatsu, M.  
 Proc. Natl. Acad. Sci. U.S.A. 79, 5443, 1982  
 A;Title: Primary structure of a cytochrome P450: coding nucleotide sequence of phenobarbital  
 A;Reference number: A93925  
 A;Accession: A93925  
 A;Note: the mistranslations in reference A93912 are acknowledged  
 R;Botelho, L.H.; Ryan, D.E.; Levin, W.  
 J. Biol. Chem. 254, 5635-5640, 1979

A;Title: Amino acid compositions and partial amino acid sequences of three highly purified or 3-methylcholanthrene.  
 A;Reference number: A32255; MUID:79194111; PMID:109438  
 A;Accession: B92255  
 A;Molecule type: protein  
 A;Residues: 1-3, 'T', 5-22 <BOT>  
 R;Lacroix, D.; Desrochers, M.; Lambert, M.; Anderson, A.  
 Gene 86, 201-207, 1990  
 A;Title: Alternative splicing of mRNA encoding rat liver cytochrome P450e (P450IIB2).  
 A;Reference number: S15589; MUID:90215299; PMID:2323573  
 A;Accession: S15589  
 A;Molecule type: mRNA  
 A;Residues: 105-113, 'F', 115-274, 'VSPAMRE', 275-321, 'E', 323-491 <IAC>  
 A;Cross-references: EMBL:M34452; NID:G203679; PIDN:AAA41004.1; PID:G203680  
 A;Note: translation of the nucleotide sequence is not complete  
 R;Phillips, I.R.; Shephard, E.A.; Ashworth, A.; Rabin, B.R.  
 Gene 24, 41-52, 1983  
 A;Reference number: A21872  
 A;Accession: A21872  
 A;Molecule type: mRNA  
 A;Residues: 168-321, 'E', 323-443, 'K', 445-491 <PHI>  
 R;Afolter, M.; Anderson, A.  
 Biochem. Biophys. Res. Commun. 118, 655-662, 1984  
 A;Title: Segmental homologies in the coding and 3' non-coding sequences of rat liver cytochrome P450  
 A;Reference number: A32736; MUID:84153837; PMID:6322758  
 A;Accession: A32736  
 A;Molecule type: mRNA  
 A;Residues: 385-491 <AFF>  
 A;Cross-references: GB:K01626; NID:G203782; PIDN:AAA41037.1; PID:G203783  
 R;Oesch, F.; Waxman, D.J.; Morrissey, J.J.; Honscha, W.; Kissel, W.; Friedberg, T.  
 Arch. Biochem. Biophys. 270, 23-32, 1989  
 A;Title: Antibodies targeted against hypervariable and constant regions of cytochromes P450  
 A;Reference number: S03854; MUID:89192373; PMID:2539047  
 A;Accession: S03854  
 A;Status: not compared with conceptual translation  
 A;Molecule type: mRNA  
 A;Residues: 329-358, 'AS', 361, 362, 363-380, 402-423 <OES>  
 R;Atchison, M.L.; Adesnik, M.  
 Proc. Natl. Acad. Sci. U.S.A. 83, 2300-2304, 1986  
 A;Title: Gene conversion in a cytochrome P-450 gene family.  
 A;Reference number: I59060; MUID:86205943; PMID:3458196  
 A;Accession: I59060  
 A;Status: preliminary; translated from GB/EMBL/DDBJ  
 A;Molecule type: DNA  
 A;Residues: 323-431 <RES>  
 A;Cross-references: GB:M13234; NID:G203848; PIDN:AAA41057.1; PID:G554434  
 C;Genetics:  
 A;Gene: CYP2B2  
 A;Introns: 384/3  
 C;Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology  
 C;Keywords: alternative splicing; chromoprotein; electron transfer; endoplasmic reticulum  
 F;295-458/Domain: cytochrome p450 homology <P45>  
 F;436/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 49.4%; Score 1293; DB 1; Length 491;  
 Best Local Similarity 50.4%; Pred. No. 2e-97;  
 Matches 243; Conservative 96; Mismatches 141; Indels 2; Gaps 2;

QY 9 LLLALLLLLLTLLSGTARGHLPGPTPLGLNLLQLRPGALYSGLMLSKYGPVF 68  
 Db 7 LLLALLVGLLLLVRGHPKSRGNFPGLPLGLNLLQLDRGLNSFMQLREKYGVF 66

QY 69 TTYLGPWRPVVVLGVQAEVREALGGQAEFFSGRGTVAMLEGTFDGHGVFFSNGERWQLR 128  
 Db 67 TVHLGF-PPVVMCGTDTKEALVQAEDFSGRGINVIEPIFKEYGVIFANGERWQLR 125

QY 129 KFTMALRDLGMKREBELIOAERCLVETFGTEGPPDPSPSLLLAQATSNVCSLLFG 188  
 Db 126 RFLSATMRDFGMGRKSVVEERIOEERCLVSELKSGALLDNTLLFHSVTNICSIVFG 185

QY 189 LRFSYEDKEQAVVRAAGGTLLGVSSGGQTYEMFSWFLRLPLPGPHKOLLHHVSTLAAPT 248  
 Db 186 BRFDYTDQFRLLELFFVTRTSLSSFSQVFEFFSGFLKYPFGAHRQIKNLQBILOYI 245



RESULT 10  
B31047  
testosterone 16alpha-hydroxylase (EC 1.14.14.14.-) cytochrome P450 2B10 - mouse  
N:Alternate names: cytochrome P450CDB; cytochrome P450p43/46  
C:Species: Mus musculus (house mouse)  
C:Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 04-Mar-2000  
C:Accession: B31047; A60559  
R:Noshiro, M.; Iakso, M.; Kawajiri, K.; Negishi, M.  
Biochemistry 27, 6434-6443, 1988  
A:Title: Rip locus: regulation of female-specific isozyme (I-P-450-16-alpha) of testosterone  
A:Reference number: A31047; MUID:89118235; PMID:3219345  
A:Accession: B31047  
A:Molecule type: mRNA  
A:Residues: 1-500 <NO>  
A:Cross-references: EMBL:M21856; NID:G201968; PIDN:AAA0425.1; PID:G201969  
A:Note: The authors translated the codon TTA for residue 55 as Phe, AGA for residue 133 as Val, AAA for residue 236 as Asn, AGG for residue 251 as Lys, AAG for residue 253 as Asp  
A:Note: The authors translated the codon GCA for residue 281 as Thr, TTC for residue 283 as Arg, ACA for residue 321 as Ala, GTG for residue 331 as Leu, TCA for residue 349 as s Ala, and GAC for residue 479 as Gly  
A:Note: The sequence nucleotide translation from Fig. 8 is inconsistent with the nucleob R:Bornheim, L.M.; Correia, M.A.  
Mol. Pharmacol. 36, 377-383, 1989  
A:Title: Purification and characterization of a mouse liver cytochrome P-450 induced by A:Reference number: A60559; MUID:89384475; PMID:2779523  
A:Accession: A60559  
A:Molecule type: protein  
A:Residues: 1-15 <BOR>  
A:Note: this enzyme was induced by cannabidiol  
C:Genetics:  
A:Gene: Cyp2b-10  
C:Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology  
C:Keywords: chromoprotein; electron transfer; endoplasmic reticulum; heme; iron; metallo F:295-467/Domain: cytochrome P450 homology <P45>  
F:445/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 49.0%; Score 1280.5; DB 2; Length 500;  
Best Local Similarity 49.7%; Pred. No. 2.1e-96;  
Matches 244; Conservative 99; Mismatches 137; Indels 11; Gaps 3;

QY 9 LLLALLALLLLTALSGTRAGHLPDGPPTPLILGNLLQLRPGALYSGLMRLSKKYGFV 68  
DB 7 LLLALLVGLFLLLARGHPSKRGNPPGPRPLILGNLLQLRPGALYSGLMRLSKKYGFV 66  
QY 69 TIVLGPWPVVLVGOEAVREALGGQAEESGRTVAMLEGTDFGHGVFSGNGRWQLR 128  
DB 67 TVHLGP-RPVVLLCGDITREALVGOEAVREALGGQAEESGRTVAMLEGTDFGHGVFSGNGRWQLR 125  
QY 129 KFTMLALRDLGMKREGEELIQAEARCLVETFOGTEGRPDPSLLLAQAATSNVVCSLFG 188  
DB 126 RFSLATMRDGMKRSVEERIQEAEQCLVEELRKSQAPLDPTFLPQCITANVCISVFG 185  
QY 189 LRFSEDEKFEQAVVRAAGTLLGVSSGGQTYEMFSLRPLPGPHKQLLHVSTLAAFT 248  
DB 186 ERFETDRLPRLLELFYQTFSLISSFSQMFELFSGFLKYPFGCAHQISKNLQELLDYI 245  
QY 249 VRQVOQHGNLDASGPARDLVADFLKMAQEEONPGTEFTKNMMLTVIYLLFAGTMTVS 308  
DB 246 GHSVERKATLDPSVP-RFIDILYLRMEKESKNQAEFHQNLMMSVLSLFPVGTETSS 304  
QY 309 TVTGYTLLLMKYPHVQKWRVEELNRELGAQAPSLGDRTRPLPYTDAVLHEAQRLLALVP 368  
DB 305 TTLHYGFLMLKYPHTEVKQEIQVIGSHRLPTLDDRTKMPYSDAVTHEIQRFSDLLIP 364  
QY 369 MGIPRLMTTRPRGYTLPGQTEVFPLLGSILHDPNIFKHPEEFNDRFLDADGRPKHE 428  
DB 365 IGVPHRVTKTLFRGYLLPKNTEVYPILSALHDPQYFEQPSFNDQFLDANGALKKSE 424  
QY 429 AFLPF-----SLGKRVCLGEGAKAELFFFTTILQAFSLSPCPDPTLSLKPVTYS 479  
DB 425 AFLPSTGQIFDOKSGVKRCICIGESITARSELFFFTSILQNFVSASHVAPKIDILTPKES 484

QY 480 GLFNIPPAFQL 490  
DB 485 GIGKIPPTYQI 495

#### RESULT 11 JT0676

Cytochrome P450 2B - green monkey  
C:Species: Cercopithecus aethiops (green monkey, grivet)  
C:Date: 18-Nov-1994 #sequence\_revision 06-Jan-1995 #text\_change 04-Mar-2000  
C:Accession: JT0676  
R:Ohmori, S.; Sakamoto, Y.; Nakasa, H.; Horie, T.; Saito, K.; Kitada, M.  
submitted to JTPID, May 1994  
A:Description: Nucleotide and amino acid sequences of monkey P450 2B gene subfamily.  
A:Reference number: JT0676  
A:Accession: JT0676  
A:Molecule type: mRNA  
A:Residues: 1-491 <OHM>  
A:Experimental source: liver  
C:Genetics:  
A:Gene: CYP2B17  
C:Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology  
C:Keywords: chromoprotein; heme; iron; metalloprotein; transmembrane protein  
F:295-458/Domain: cytochrome P450 homology <P45>  
F:436/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 48.8%; Score 1275; DB 2; Length 491;  
Best Local Similarity 49.3%; Pred. No. 5.8e-96;  
Matches 241; Conservative 100; Mismatches 144; Indels 4; Gaps 3;

QY 8 ALLIALALL--LLLTALSGTRAGHLPDGPPTPLILGNLLQLRPGALYSGLMRLSKKYG 65  
DB 4 SVLLFTALLTGLLLLVQRHFNHAGRLPPGCPPLILGNLLQMRDRGLLRFLRFRKYG 63  
QY 66 PVFTYLGPMRPVVVLVGOEAVREALGGQAEESGRTVAMLEGTDFGHGVFSGNGRW 125  
DB 64 DVTVLGP-RPVVLLCGVEAREALVDNAEASGSGKIAITDVPVFGYGVIFANGNRWK 122  
QY 126 QLRFKTMALRDLGMKREGEELIQAEARCLVETFOGTEGRPDPSLLLAQAATSNVVCSL 185  
DB 123 VLARFSLTTRDGMKRSVEERIQEAEQCLVEELRKSQALVDPTFLFHSITANIICSI 182  
QY 186 LFGIRSYEDKEFEQAVVRAAGTLLGVSSGGQTYEMFSLRPLPGPHKQLLHVSTLA 245  
DB 183 VFGRHYQDEFLKINLFYHTFSLASSMFGQLFELLGFLKYPFGCAHQVYKNLQEI 242  
QY 246 AFTVROVQOQHGNLDASGPARDLVADFLKMAQEEONPGTEFTKNMMLTVIYLLFAGTM 305  
DB 243 AYIGHSVKREHRETLDPAP-QDLIDSYLQMEKESKNPHSEFHSRNLIIINTLSLFFAGTE 301  
QY 306 TVSTTVGYTLLLMKYPHVQKWRVEELNRELGAQAPSLGDRTRPLPYTDAVLHEAQRLLA 365  
DB 302 TSTTLRYGFLMLKYPHVAERYKEIQVIGPHRPPALDDRAKMPYTEAVTHEIQRFAD 361  
QY 365 LVPMGIPRLMTTRPRGYTLPGQTEVFPLLGSILHDPNIFKHPEEFNDRFLDADGRFR 425  
DB 362 LUPMGVPHVITQOTSERGYIIPKDTVEVFLSTALHDPHYFEKPDFTFNDPDLANGALK 421  
QY 426 KHEAFPFSLGKRVCLGEGAKAELFFFTTILQAFSLSPCPDPTLSLKPVTYSGLFNIP 485  
DB 422 KNEAFIPFSLGRMCLGEGIARENLFFFTTILQNFVSASVPAPEDIDLTPQESGVGKIP 481  
QY 486 PAFOQVRP 494  
DB 482 PTYQIRFLP 490

#### RESULT 12 S11305

Cytochrome P450 2B11 - dog  
N:Alternate names: Cytochrome P450IIB  
N:Contains: oxidoreductase [EC 1.-.-.-]  
C:Species: Canis lupus familiaris (dog)



Query Match 47.6%; Score 1244; DB 2; Length 491;  
Best Local Similarity 48.5%; Pred. No. 1.9e-93;  
Matches 234; Conservative 100; Mismatches 146; Indels 2; Gaps 2;

QY 9 LLLALLLLLTTALSQTRARGHLPGGTPLLGNLQLRPGALYSGLMRLSKYGPVF 68  
DB :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:  
7 LLLAVLSLFLLVRGHAKIHGHPGPHPLPGLLNLMQDRGGLLKCFIQEKHGDFV 66  
DB :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:  
QY 69 TIYLGPWRPVVLVGQEAVRREALGQAEBFSGRGTVAMLEGTFDGHGVFFSNGERWROLR 128  
DB :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:  
67 TVHLGP-RPVVVLCGTQTIREALVDHAFAFSGRGTIAAAQLVMQDYGIFFASGQRWLIR 125  
DB :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:  
QY 129 KFTMLALRDLCWGKEEGEELTOAERARCLVEFQGTGEPFPDPSSLILAQATSNVVCSLFG 188  
DB :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:  
126 RFLSATWKFEFGMKRSVEERIKEEAQCLVELKKYQGVPFLDPTFLFCQCITANIICSIIVFG 185  
DB :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:  
QY 189 LRFSYEDKFEQAVRAAGCTLLGVSSQCGQTVEMPSWFLRPLRPGPKOLLHHVSTLAAPT 248  
DB :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:  
186 ERFDYTDQFLHLNLMVKIFSLSSFSGQMFELPSGFLKYFPGVHRQIVVKKQELLDTI 245  
DB :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:  
QY 249 VRVQQHQGNLDASGPARDLVDAFLFKWAQEONGPTEFTNNKNMLMTVIYLLFACTMTVS 308  
DB :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:  
246 AHSVEKHATLDPSAP-RDYIDTYLLRMKEKSNHNTEFHQNLMMSVLSFFVGTETTS 304  
DB :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:  
QY 309 TTVGYTLLLMKYPHVQKWVREELNRRELGAQAOSLGDRTPLYTDAVLHEAORLLALVP 368  
DB :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:  
305 ATLHYGVLLMKYPHTVEKVQKEIDQVTIGSHRLPTLDDRTKMPYTDVAIHIQRFSDIJP 364  
DB :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:  
QY 369 MGIPRTLMTTRFRGYTLPOQTEVPFLLGSILHDPNFIKHPPEENPORFLDADGRFKHE 428  
DB :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:  
365 IGLPHKVIKOTLFRGYLLPKNTENVYVVLSSALHDPQYFEQDKFNPEQFLDANGALKCE 424  
DB :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:  
QY 429 AFUPESIGKVCVCEGLAKAELEFPFTTILOAFSLESPPCPDPTLSLKPTVSGLENIPPAF 488  
DB :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:  
425 AFUPFSTGKRKICGESIARNELFFFTTILQNFVSASPAPKDIDLTPKESGICKIPPAH 484  
DB :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:  
QY 489 QL 490  
DB :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:  
485 QI 486

RESULT 15  
A3551  
cytochrome P450 2G1 - rat  
N:Contains: oxidoreductase (EC 1.-.-.-)  
C:Species: Rattus norvegicus (Norway rat)  
C>Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 04-Mar-2000  
C:Accession: A35551; A33875  
Ri:Nef, P.; Larabee, T.M.; Kagimoto, K.; Meyer, U.A.  
J. Biol. Chem. 265, 2903-2907, 1990  
A>Title: Olfactory-specific cytochrome P-450 (P-450olf; IIGl). Gene structure and devel  
A:Reference number: A35551; MUID:90153922; PMID:2406242  
A:Accession: A35551  
A:Molecule type: DNA  
A:Residues: 1-494 <NEF>  
A:Cross-references: GB:M31931; GB:J04715; NID:g203888; PIDN:AAAA1069.1; PID:g203890  
Ri:Nef, P.; Heldman, J.; Lazard, D.; Margalit, T.; Jaye, M.; Hanukoglu, I.; Lancet, D.  
J. Biol. Chem. 264, 6780-6785, 1989  
A>Title: Olfactory-specific cytochrome P-450. cDNA cloning of a novel neuroepithelial en  
A:Reference number: A33875; MUID:89214086; PMID:2708343  
A:Accession: A33875  
A:Molecule type: mRNA  
A:Residues: 2-494 <NE2>  
A:Cross-references: GB:M33296; GB:M31931; GB:J04715; NID:g203888; PIDN:AAAA1069.1; PID:g  
C:Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology  
C:Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; monooxygenase;  
F:298-461/Domain: cytochrome P450 homology <P45>  
F:439/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 47.6%; Score 1243.5; DB 2; Length 494;  
Best Local Similarity 47.7%; Pred. No. 2.1e-93;  
Matches 234; Conservative 101; Mismatches 153; Indels 3; Gaps 3;

*This Page Blank (uspto)*



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

# OM protein - protein search, using sw model

Run on: September 15, 2004, 09:21:16 ; Search time 128 Seconds  
(without alignments)  
1262.715 Million cell updates/sec

Title: US-10-669-693-2

Perfect score: 2615

Sequence: 1 MEATGTWALLALALLLLT.....PPAFQLQVRPTDLHSTTQTR 504

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1335176 seqs, 320689617 residues

Total number of hits satisfying chosen parameters: 1335176

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

## Database :

Published Applications AA:\*  
1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	% Match	Query Match	Length	ID	Description
1	2615	100.0	504	9	US-09-748-127-2	Sequence 2, Appli
2	2615	100.0	504	12	US-10-147-493-64	Sequence 64, Appli
3	2615	100.0	504	12	US-10-258-080-6	Sequence 6, Appli
4	2615	100.0	504	12	US-10-145-127-64	Sequence 64, Appli
5	2615	100.0	504	12	US-10-160-503-64	Sequence 64, Appli
6	2615	100.0	504	12	US-10-143-118-64	Sequence 64, Appli
7	2615	100.0	504	12	US-10-144-993-64	Sequence 64, Appli
8	2615	100.0	504	12	US-10-158-787-64	Sequence 64, Appli
9	2615	100.0	504	12	US-10-669-693-2	Sequence 2, Appli
10	2615	100.0	504	12	US-10-140-024-64	Sequence 64, Appli
11	2615	100.0	504	12	US-10-140-808-64	Sequence 64, Appli
12	2615	100.0	504	12	US-10-152-405-64	Sequence 64, Appli
13	2615	100.0	504	12	US-10-127-852A-64	Sequence 64, Appli
14	2615	100.0	504	12	US-10-127-900A-64	Sequence 64, Appli
15	2615	100.0	504	12	US-10-128-685A-64	Sequence 64, Appli

16	2615	100.0	504	12	US-10-131-820A-64	Sequence 64, Appli
17	2615	100.0	504	12	US-10-142-886-64	Sequence 64, Appli
18	2615	100.0	504	12	US-10-146-728-64	Sequence 64, Appli
19	2615	100.0	504	12	US-10-146-786-64	Sequence 64, Appli
20	2615	100.0	504	12	US-10-147-499-64	Sequence 64, Appli
21	2615	100.0	504	12	US-10-157-798-64	Sequence 64, Appli
22	2615	100.0	504	14	US-10-028-072-64	Sequence 5, Appli
23	2615	100.0	504	14	US-10-121-049-64	Sequence 64, Appli
24	2615	100.0	504	14	US-10-123-904-64	Sequence 64, Appli
25	2615	100.0	504	14	US-10-140-470-64	Sequence 64, Appli
26	2615	100.0	504	14	US-10-067-668-5	Sequence 64, Appli
27	2615	100.0	504	14	US-10-175-746-64	Sequence 5, Appli
28	2615	100.0	504	14	US-10-176-918-64	Sequence 64, Appli
29	2615	100.0	504	14	US-10-176-921-64	Sequence 64, Appli
30	2615	100.0	504	14	US-10-137-865-64	Sequence 64, Appli
31	2615	100.0	504	14	US-10-140-474-64	Sequence 64, Appli
32	2615	100.0	504	14	US-10-142-431-64	Sequence 64, Appli
33	2615	100.0	504	14	US-10-143-114-64	Sequence 64, Appli
34	2615	100.0	504	14	US-10-140-002-64	Sequence 64, Appli
35	2615	100.0	504	14	US-10-142-419-64	Sequence 64, Appli
36	2615	100.0	504	14	US-10-123-262-64	Sequence 64, Appli
37	2615	100.0	504	14	US-10-142-423-64	Sequence 64, Appli
38	2615	100.0	504	14	US-10-121-050-64	Sequence 64, Appli
39	2615	100.0	504	14	US-10-141-755-64	Sequence 64, Appli
40	2615	100.0	504	14	US-10-143-032-64	Sequence 64, Appli
41	2615	100.0	504	14	US-10-123-108-64	Sequence 64, Appli
42	2615	100.0	504	14	US-10-123-236-64	Sequence 64, Appli
43	2615	100.0	504	14	US-10-123-261-64	Sequence 64, Appli
44	2615	100.0	504	14	US-10-140-921-64	Sequence 64, Appli
45	2615	100.0	504	14	US-10-140-928-64	Sequence 64, Appli

## ALIGNMENTS

RESULT 1  
US-09-748-127-2  
; Sequence 2, Application US/09748127  
; Patent No. US2002007674A1  
; GENERAL INFORMATION:  
; APPLICANT: YAN, Chunhua et al  
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING  
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN  
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS,  
; TITLE OF INVENTION: AND USES THEREOF  
; FILE REFERENCE: CL000685  
; CURRENT APPLICATION NUMBER: US/09/748,127  
; CURRENT FILING DATE: 2001-06-11  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 504  
; TYPE: PRT  
; ORGANISM: Human  
US-09-748-127-2

Query Match	100.0%;	Score 2615;	DB 9;	Length 504;
Best Local Similarity	100.0%;	Pred. No. 8.1e-253;		
Matches 504;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MEATGTWALLALALLLLTALSGTRARGHLP	PPGTPPLPGLGNLLQLRFGALYGLMRL	60
Db	1	MEATGTWALLALALLLLTALSGTRARGHLP	PPGTPPLPGLGNLLQLRFGALYGLMRL	60
QY	61	SKKYGVFTYILGPNRPVVVLVGQAVREALGQAE	EEFSGRGTVMLEGTDFGHGVFFSN	120
Db	61	SKKYGVFTYILGPNRPVVVLVGQAVREALGQAE	EEFSGRGTVMLEGTDFGHGVFFSN	120
QY	121	GERWQLRKFTMLALDLGMKREGEELIQAEARCL	VETFGTEGPRPPPSLLLAQATSN	180
Db	121	GERWQLRKFTMLALDLGMKREGEELIQAEARCL	VETFGTEGPRPPPSLLLAQATSN	180
QY	181	VVCSLLFLGRFSYEDKEFQAVVRAAGTLLGVSS	QGGQGTVMFSLRPLPGPKQLLHH	240

```
Db      121  GERWQLRKFTMLALRDLGMGKREGEELIQAEARCLVETFGTEGRPFDPSPSLLLAQATSN 180
Qy      181  VVCSLLFGLRSYEDKQFQAVVRAAGGTLGVSQGGQTYEMFQWFLRPLPQPHKQLLHH 240
Db      181  VVCSLLFGLRSYEDKQFQAVVRAAGGTLGVSQGGQTYEMFQWFLRPLPQPHKQLLHH 240
Qy      241  VSTLAAFTVRQVQOQGNLDASGPARDLDVDAFLKMAQEQNPGTEFTNKMLMTVIYLL 300
Db      241  VSTLAAFTVRQVQOQGNLDASGPARDLDVDAFLKMAQEQNPGTEFTNKMLMTVIYLL 300
Qy      301  FAGTMVSTTVGYTLLLMKYPHVQKVRREELNRELGLGAGQAPSLGDRTRLPYTDVHLHA 360
Db      301  FAGTMVSTTVGYTLLLMKYPHVQKVRREELNRELGLGAGQAPSLGDRTRLPYTDVHLHA 360
Qy      361  ORLLALVPMGIPRTLMRTTRFRGYTLPGQTEVPFLGSLILHDPNFKHPEEENPDRFLDA 420
Db      361  ORLLALVPMGIPRTLMRTTRFRGYTLPGQTEVPFLGSLILHDPNFKHPEEENPDRFLDA 420
Qy      421  DGRFRKHEAFLPFSLGKRVCLGEGLAKEALFLFFTTILQAFSLESFPCPPDITLSLKPTVSG 480
Db      421  DGRFRKHEAFLPFSLGKRVCLGEGLAKEALFLFFTTILQAFSLESFPCPPDITLSLKPTVSG 480
Qy      481  LFNIPPAFQQLVRPTDLHSTTQTR 504
Db      481  LFNIPPAFQQLVRPTDLHSTTQTR 504

RESULT 3
US-10-258-080-6
; Sequence 6, Application US/10258080
; Publication No. US20040029125A1
; GENERAL INFORMATION:
; APPLICANT: Incyte Genomics, Inc.
; APPLICANT: POLICKY, Jennifer L.
; APPLICANT: HAFALIA, April J.A.
; APPLICANT: BURFORD, Neil
; APPLICANT: RING, Huijun Z.
; APPLICANT: LAL, Preeti G.
; APPLICANT: TRIBOULEY, Catherine M.
; APPLICANT: YAO, Monique G.
; APPLICANT: TANG, Y. Tom
; APPLICANT: ARVIZU, Chandra S.
; APPLICANT: DAS, Debopriya
; APPLICANT: SANJANWALA, Madhusudan M.
; APPLICANT: GANDHI, Ameena R.
; APPLICANT: REDDY, Roopa M.
; APPLICANT: KHAN, Farrah A.
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: RAMKUMAR, Jayalaxmi
; APPLICANT: GRIFFIN, Jennifer A.
; APPLICANT: AU-YOUNG, Janice K.
; TITLE OF INVENTION: DRUG METABOLIZING ENZYMES
; FILE REFERENCE: PI-0070 USN
; CURRENT APPLICATION NUMBER: US/10/258,080
; CURRENT FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: US 60/203,509
; PRIOR FILING DATE: 2000-05-11
; PRIOR APPLICATION NUMBER: US 60/202,234
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/200,185
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/198,403
; PRIOR FILING DATE: 2000-04-19
; PRIOR APPLICATION NUMBER: PCT/US01/11869
; PRIOR FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: US 60/197,590
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PERL Program
; SEQ ID NO 6
; LENGTH: 504
; TYPE: PRT
```

```
Db      181  VVCSLLFGLRSYEDKQFQAVVRAAGGTLGVSQGGQTYEMFQWFLRPLPQPHKQLLHH 240
Qy      241  VSTLAAFTVRQVQOQGNLDASGPARDLDVDAFLKMAQEQNPGTEFTNKMLMTVIYLL 300
Db      241  VSTLAAFTVRQVQOQGNLDASGPARDLDVDAFLKMAQEQNPGTEFTNKMLMTVIYLL 300
Qy      301  FAGTMVSTTVGYTLLLMKYPHVQKVRREELNRELGLGAGQAPSLGDRTRLPYTDVHLHA 360
Db      301  FAGTMVSTTVGYTLLLMKYPHVQKVRREELNRELGLGAGQAPSLGDRTRLPYTDVHLHA 360
Qy      361  ORLLALVPMGIPRTLMRTTRFRGYTLPGQTEVPFLGSLILHDPNFKHPEEENPDRFLDA 420
Db      361  ORLLALVPMGIPRTLMRTTRFRGYTLPGQTEVPFLGSLILHDPNFKHPEEENPDRFLDA 420
Qy      421  DGRFRKHEAFLPFSLGKRVCLGEGLAKEALFLFFTTILQAFSLESFPCPPDITLSLKPTVSG 480
Db      421  DGRFRKHEAFLPFSLGKRVCLGEGLAKEALFLFFTTILQAFSLESFPCPPDITLSLKPTVSG 480
Qy      481  LFNIPPAFQQLVRPTDLHSTTQTR 504
Db      481  LFNIPPAFQQLVRPTDLHSTTQTR 504

RESULT 2
US-10-147-493-64
; Sequence 64, Application US/10147493
; Publication No. US20040029217A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: F330R1C345
; CURRENT APPLICATION NUMBER: US/10/147,493
; CURRENT FILING DATE: 2002-05-17
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 64
; LENGTH: 504
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-147-493-64

Query Match      100.0%; Score 2615; DB 12; Length 504;
Best Local Similarity 100.0%; Pred. No. 8.1e-253;
Matches 504; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1  MEATGTWALLLALLLTLALSGTRAGHLPGPPTPLPLGNLLQLRPGALYSGMLRL 60
Db      1  MEATGTWALLLALLLTLALSGTRAGHLPGPPTPLPLGNLLQLRPGALYSGMLRL 60
Qy      61  SKYGPVFTTYLGPWRPVWLVGQAVREALGGQAEFFSGRTVAMLEGTFDGHGVFFSN 120
Db      61  SKYGPVFTTYLGPWRPVWLVGQAVREALGGQAEFFSGRTVAMLEGTFDGHGVFFSN 120
Qy      121  GERWQLRKFTMLALRDLGMGKREGEELIQAEARCLVETFGTEGRPFDPSPSLLLAQATSN 180
```

```
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20040029125A1 1272843CD1
US-10-258-080-6

Query Match      100.0%; Score 2615; DB 12; Length 504;
Best Local Similarity 100.0%; Pred. No. 8.1e-253;
Matches 504; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEATGTWALLALALLLLTALSGTRARGHLPPGPTPLPLGNLLQLRPGALYSGLMRL 60
Db 1 MEATGTWALLALALLLLTALSGTRARGHLPPGPTPLPLGNLLQLRPGALYSGLMRL 60
QY 61 SKYGVFTIYLGPNRPVVVLVGOEAVREALGQAEFFSGRGTVAMLEGTDFGHGVFFSN 120
Db 61 SKYGVFTIYLGPNRPVVVLVGOEAVREALGQAEFFSGRGTVAMLEGTDFGHGVFFSN 120
QY 121 GERWRQLRKFTMLALRDGLMGKEGEBELIOAEARCLVETFGTEGPRPDPSSLQAQTSN 180
Db 121 GERWRQLRKFTMLALRDGLMGKEGEBELIOAEARCLVETFGTEGPRPDPSSLQAQTSN 180
QY 181 VVCSLLFGLRFSYEDKEFOAVVRAAGTLLGVSSQGGQTYEMFSWFLRPLPGPHKQLLHH 240
Db 181 VVCSLLFGLRFSYEDKEFOAVVRAAGTLLGVSSQGGQTYEMFSWFLRPLPGPHKQLLHH 240
QY 241 VSTLAFTVRQVOHQGNLDASGPARDLVDAFLLKMAQEEQNPGTEFTNKNMLMTVIYLL 300
Db 241 VSTLAFTVRQVOHQGNLDASGPARDLVDAFLLKMAQEEQNPGTEFTNKNMLMTVIYLL 300
QY 301 FACTMTVSTTVGTYLLMLMKYPHVQKWVREBELNRELGAQAPSLGDRTRLPYTDVILHEA 360
Db 301 FACTMTVSTTVGTYLLMLMKYPHVQKWVREBELNRELGAQAPSLGDRTRLPYTDVILHEA 360
QY 361 QRLALVPNGIPRTLMRTTRFRGYTLPGQTEVPFLGSIILHDPNIFKHPEEFPDRFLDA 420
Db 361 QRLALVPNGIPRTLMRTTRFRGYTLPGQTEVPFLGSIILHDPNIFKHPEEFPDRFLDA 420
QY 421 DGRFRKHEAPFLPSLGRKVCGLGELAKAEFLFPTTILQAFSLESPPDPTLSLKPTVSG 480
Db 421 DGRFRKHEAPFLPSLGRKVCGLGELAKAEFLFPTTILQAFSLESPPDPTLSLKPTVSG 480
QY 481 LFNIPPAFOLQVRPTDLHSTTQTR 504
Db 481 LFNIPPAFOLQVRPTDLHSTTQTR 504

RESULT 4
US-10-145-127-64
; Sequence 64, Application US/10145127
; Publication No. US20040033558A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C252
; CURRENT APPLICATION NUMBER: US/10/145,127
```

```
; CURRENT FILING DATE: 2002-05-13
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 64
; LENGTH: 504
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-145-127-64

Query Match      100.0%; Score 2615; DB 12; Length 504;
Best Local Similarity 100.0%; Pred. No. 8.1e-253;
Matches 504; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEATGTWALLALALLLLTALSGTRARGHLPPGPTPLPLGNLLQLRPGALYSGLMRL 60
Db 1 MEATGTWALLALALLLLTALSGTRARGHLPPGPTPLPLGNLLQLRPGALYSGLMRL 60
QY 61 SKYGVFTIYLGPNRPVVVLVGOEAVREALGQAEFFSGRGTVAMLEGTDFGHGVFFSN 120
Db 61 SKYGVFTIYLGPNRPVVVLVGOEAVREALGQAEFFSGRGTVAMLEGTDFGHGVFFSN 120
QY 121 GERWRQLRKFTMLALRDGLMGKEGEBELIOAEARCLVETFGTEGPRPDPSSLQAQTSN 180
Db 121 GERWRQLRKFTMLALRDGLMGKEGEBELIOAEARCLVETFGTEGPRPDPSSLQAQTSN 180
QY 181 VVCSLLFGLRFSYEDKEFOAVVRAAGTLLGVSSQGGQTYEMFSWFLRPLPGPHKQLLHH 240
Db 181 VVCSLLFGLRFSYEDKEFOAVVRAAGTLLGVSSQGGQTYEMFSWFLRPLPGPHKQLLHH 240
QY 241 VSTLAFTVRQVOHQGNLDASGPARDLVDAFLLKMAQEEQNPGTEFTNKNMLMTVIYLL 300
Db 241 VSTLAFTVRQVOHQGNLDASGPARDLVDAFLLKMAQEEQNPGTEFTNKNMLMTVIYLL 300
QY 301 FACTMTVSTTVGTYLLMLMKYPHVQKWVREBELNRELGAQAPSLGDRTRLPYTDVILHEA 360
Db 301 FACTMTVSTTVGTYLLMLMKYPHVQKWVREBELNRELGAQAPSLGDRTRLPYTDVILHEA 360
QY 361 QRLALVPNGIPRTLMRTTRFRGYTLPGQTEVPFLGSIILHDPNIFKHPEEFPDRFLDA 420
Db 361 QRLALVPNGIPRTLMRTTRFRGYTLPGQTEVPFLGSIILHDPNIFKHPEEFPDRFLDA 420
QY 421 DGRFRKHEAPFLPSLGRKVCGLGELAKAEFLFPTTILQAFSLESPPDPTLSLKPTVSG 480
Db 421 DGRFRKHEAPFLPSLGRKVCGLGELAKAEFLFPTTILQAFSLESPPDPTLSLKPTVSG 480
QY 481 LFNIPPAFOLQVRPTDLHSTTQTR 504
Db 481 LFNIPPAFOLQVRPTDLHSTTQTR 504

RESULT 5
US-10-160-503-64
; Sequence 64, Application US/10160503
; Publication No. US20040033559A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
```

```
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C446
; CURRENT APPLICATION NUMBER: US/10/160,503
; CURRENT FILING DATE: 2002-05-30
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 64
; LENGTH: 504
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-160-503-64

Query Match      100.0%; Score 2615; DB 12; Length 504;
Best Local Similarity 100.0%; Pred. No. 8.1e-253;
Matches 504; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEATGTWALLALALLLLLTALSGTRARGHLPDGPPTPLPLGNLLQLRPGALYGLMRL 60
DB 1 MEATGTWALLALALLLLLTALSGTRARGHLPDGPPTPLPLGNLLQLRPGALYGLMRL 60
QY 61 SKKYGPVFTIYLGPRPVPVVLVGQAVREALGQAEFSGRGTVAMLEGTDFDGHGVFFSN 120
DB 61 SKKYGPVFTIYLGPRPVPVVLVGQAVREALGQAEFSGRGTVAMLEGTDFDGHGVFFSN 120
QY 121 GERWQLRKFTMLALRDLGMKREGEELIQAEARCLVETFOGTEGRPFDPSSLILAQATSN 180
DB 121 GERWQLRKFTMLALRDLGMKREGEELIQAEARCLVETFOGTEGRPFDPSSLILAQATSN 180
QY 181 VVCSLLFGLRFSYEDKEFOAVVRAAGTLLGVSSQGGQTYEMFSLRPLPGPHKQLLHH 240
DB 181 VVCSLLFGLRFSYEDKEFOAVVRAAGTLLGVSSQGGQTYEMFSLRPLPGPHKQLLHH 240
QY 241 VSTLAAFTVRQVQOQHGNDASGPARDLVDAFLLMKMAQEEQNPGETFNKNMLMTVIYLL 300
DB 241 VSTLAAFTVRQVQOQHGNDASGPARDLVDAFLLMKMAQEEQNPGETFNKNMLMTVIYLL 300
QY 301 FAGTMTVSTTVGYTLLMLKYPHVQKVRBELNRELGLGAGQPSLGDRTLRPYTDAVLHEA 360
DB 301 FAGTMTVSTTVGYTLLMLKYPHVQKVRBELNRELGLGAGQPSLGDRTLRPYTDAVLHEA 360
QY 361 QRLALALVPMGIPRTLMTTRFRGYTLPGQTEVFPPLGSLILHDPNIFKHIPEENPDRFLDA 420
DB 361 QRLALALVPMGIPRTLMTTRFRGYTLPGQTEVFPPLGSLILHDPNIFKHIPEENPDRFLDA 420
QY 421 DGRFRKHEAFLPFSLGKRVCLGEGLAELFLFFTTILQAFLSLESPCPDPTLSLKPTVSG 480
DB 421 DGRFRKHEAFLPFSLGKRVCLGEGLAELFLFFTTILQAFLSLESPCPDPTLSLKPTVSG 480
QY 481 LFNIPPAFQLQVRPTDLHSTTTQTR 504
DB 481 LFNIPPAFQLQVRPTDLHSTTTQTR 504

RESULT 6
US-10-143-118-64
; Sequence 64, Application US/10143118
; Publication No. US20040038335A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tamas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C228
; CURRENT APPLICATION NUMBER: US/10/143,118
; CURRENT FILING DATE: 2002-05-09
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 64
; LENGTH: 504
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-143-118-64

Query Match      100.0%; Score 2615; DB 12; Length 504;
Best Local Similarity 100.0%; Pred. No. 8.1e-253;
Matches 504; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEATGTWALLALALLLLLTALSGTRARGHLPDGPPTPLPLGNLLQLRPGALYGLMRL 60
DB 1 MEATGTWALLALALLLLLTALSGTRARGHLPDGPPTPLPLGNLLQLRPGALYGLMRL 60
QY 61 SKKYGPVFTIYLGPRPVPVVLVGQAVREALGQAEFSGRGTVAMLEGTDFDGHGVFFSN 120
DB 61 SKKYGPVFTIYLGPRPVPVVLVGQAVREALGQAEFSGRGTVAMLEGTDFDGHGVFFSN 120
QY 121 GERWQLRKFTMLALRDLGMKREGEELIQAEARCLVETFOGTEGRPFDPSSLILAQATSN 180
DB 121 GERWQLRKFTMLALRDLGMKREGEELIQAEARCLVETFOGTEGRPFDPSSLILAQATSN 180
QY 181 VVCSLLFGLRFSYEDKEFOAVVRAAGTLLGVSSQGGQTYEMFSLRPLPGPHKQLLHH 240
DB 181 VVCSLLFGLRFSYEDKEFOAVVRAAGTLLGVSSQGGQTYEMFSLRPLPGPHKQLLHH 240
QY 241 VSTLAAFTVRQVQOQHGNDASGPARDLVDAFLLMKMAQEEQNPGETFNKNMLMTVIYLL 300
DB 241 VSTLAAFTVRQVQOQHGNDASGPARDLVDAFLLMKMAQEEQNPGETFNKNMLMTVIYLL 300
QY 301 FAGTMTVSTTVGYTLLMLKYPHVQKVRBELNRELGLGAGQPSLGDRTLRPYTDAVLHEA 360
DB 301 FAGTMTVSTTVGYTLLMLKYPHVQKVRBELNRELGLGAGQPSLGDRTLRPYTDAVLHEA 360
QY 361 QRLALALVPMGIPRTLMTTRFRGYTLPGQTEVFPPLGSLILHDPNIFKHIPEENPDRFLDA 420
DB 361 QRLALALVPMGIPRTLMTTRFRGYTLPGQTEVFPPLGSLILHDPNIFKHIPEENPDRFLDA 420
QY 421 DGRFRKHEAFLPFSLGKRVCLGEGLAELFLFFTTILQAFLSLESPCPDPTLSLKPTVSG 480
DB 421 DGRFRKHEAFLPFSLGKRVCLGEGLAELFLFFTTILQAFLSLESPCPDPTLSLKPTVSG 480
QY 481 LFNIPPAFQLQVRPTDLHSTTTQTR 504
DB 481 LFNIPPAFQLQVRPTDLHSTTTQTR 504

RESULT 7
US-10-144-993-64
; Sequence 64, Application US/10144993
; Publication No. US20040038336A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
```

; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3330R1C261  
; CURRENT APPLICATION NUMBER: US/10/144,993  
; Prior Application removed - See File Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 550  
; SEQ ID NO 64  
; LENGTH: 504  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-144-993-64

Query Match 100.0%; Score 2615; DB 12; Length 504;  
Best Local Similarity 100.0%; Pred. No. 8.1e-253;  
Matches 504; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MEATGTWALLALALLLLTALSGTRAGHLPGPTPLPLGNLLQRLPGALYSGLMRL 60  
Db 1 MEATGTWALLALALLLLTALSGTRAGHLPGPTPLPLGNLLQRLPGALYSGLMRL 60  
QY 61 SKKYGVPVFTIYLGPRPVVVLVGQAVREALGGQAEFFSGRTVAMLEGTDFDGHGVFFSN 120  
Db 61 SKKYGVPVFTIYLGPRPVVVLVGQAVREALGGQAEFFSGRTVAMLEGTDFDGHGVFFSN 120  
QY 121 GERWQLRKFTMLALRDLMGKREGELIQAEARCLVETFGTEGRPDPSLLAQAATSN 180  
Db 121 GERWQLRKFTMLALRDLMGKREGELIQAEARCLVETFGTEGRPDPSLLAQAATSN 180  
QY 181 VVCSLLFGLRFSYEDKEFQAVVRAAGTLLGVSSQGGQYEMFSLRPLPGPHKQLLHH 240  
Db 181 VVCSLLFGLRFSYEDKEFQAVVRAAGTLLGVSSQGGQYEMFSLRPLPGPHKQLLHH 240  
QY 241 VSTLAAFTVRQVQOQHGNLDASGPARDLVDAFLKMAQEEQNPGTEFTNKMLMTVIYLL 300  
Db 241 VSTLAAFTVRQVQOQHGNLDASGPARDLVDAFLKMAQEEQNPGTEFTNKMLMTVIYLL 300  
QY 301 FAGTMTVSTTVGYTLLLMKYPHVQKWVREELNRELGAGQAPSLGDRTRLPYTDVHLHEA 360  
Db 301 FAGTMTVSTTVGYTLLLMKYPHVQKWVREELNRELGAGQAPSLGDRTRLPYTDVHLHEA 360  
QY 361 QRLALVPMGIPRTLMRTTRFRGYTLPGQTEVPFLGSLIHDNIFKHPEEFNPDRLDA 420  
Db 361 QRLALVPMGIPRTLMRTTRFRGYTLPGQTEVPFLGSLIHDNIFKHPEEFNPDRLDA 420  
QY 421 DGRFRKHEAFLPFLSKRVCLGEGAKAEFLFPTTILQAFSLBSPCPDPTLSLKPTVSG 480  
Db 421 DGRFRKHEAFLPFLSKRVCLGEGAKAEFLFPTTILQAFSLBSPCPDPTLSLKPTVSG 480  
QY 481 LFNIPAFQLQVRPTDLHSTTQTR 504  
Db 481 LFNIPAFQLQVRPTDLHSTTQTR 504

## RESULT 8

US-10-158-787-64  
; Sequence 64, Application US/10158787  
; Publication No. US20040039164A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: DeForge, Laura  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.

; APPLICANT: Sherwood, Steven  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3330R1C449  
; CURRENT APPLICATION NUMBER: US/10/158,787  
; CURRENT FILING DATE: 2003-04-03  
; PRIOR APPLICATION NUMBER: 60/049911  
; PRIOR FILING DATE: 1997-06-18  
; PRIOR APPLICATION NUMBER: 60/056974  
; PRIOR FILING DATE: 1997-08-26  
; PRIOR APPLICATION NUMBER: 60/059113  
; PRIOR FILING DATE: 1997-09-17  
; PRIOR APPLICATION NUMBER: 60/059115  
; PRIOR FILING DATE: 1997-09-17  
; PRIOR APPLICATION NUMBER: 60/059117  
; PRIOR FILING DATE: 1997-09-17  
; PRIOR APPLICATION NUMBER: 60/059122  
; PRIOR FILING DATE: 1997-09-17  
; PRIOR APPLICATION NUMBER: 60/059184  
; PRIOR FILING DATE: 1997-09-17  
; PRIOR APPLICATION NUMBER: 60/059263  
; PRIOR FILING DATE: 1997-09-18  
; PRIOR APPLICATION NUMBER: 60/059352  
; PRIOR FILING DATE: 1997-09-19  
; PRIOR APPLICATION NUMBER: 60/059588  
; PRIOR FILING DATE: 1997-09-19  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 550  
; SEQ ID NO 64  
; LENGTH: 504  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-158-787-64

Query Match 100.0%; Score 2615; DB 12; Length 504;  
Best Local Similarity 100.0%; Pred. No. 8.1e-253;  
Matches 504; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MEATGTWALLALALLLLTALSGTRAGHLPGPTPLPLGNLLQRLPGALYSGLMRL 60  
Db 1 MEATGTWALLALALLLLTALSGTRAGHLPGPTPLPLGNLLQRLPGALYSGLMRL 60  
QY 61 SKKYGVPVFTIYLGPRPVVVLVGQAVREALGGQAEFFSGRTVAMLEGTDFDGHGVFFSN 120  
Db 61 SKKYGVPVFTIYLGPRPVVVLVGQAVREALGGQAEFFSGRTVAMLEGTDFDGHGVFFSN 120  
QY 121 GERWQLRKFTMLALRDLMGKREGELIQAEARCLVETFGTEGRPDPSLLAQAATSN 180  
Db 121 GERWQLRKFTMLALRDLMGKREGELIQAEARCLVETFGTEGRPDPSLLAQAATSN 180  
QY 181 VVCSLLFGLRFSYEDKEFQAVVRAAGTLLGVSSQGGQYEMFSLRPLPGPHKQLLHH 240  
Db 181 VVCSLLFGLRFSYEDKEFQAVVRAAGTLLGVSSQGGQYEMFSLRPLPGPHKQLLHH 240  
QY 241 VSTLAAFTVRQVQOQHGNLDASGPARDLVDAFLKMAQEEQNPGTEFTNKMLMTVIYLL 300  
Db 241 VSTLAAFTVRQVQOQHGNLDASGPARDLVDAFLKMAQEEQNPGTEFTNKMLMTVIYLL 300  
QY 301 FAGTMTVSTTVGYTLLLMKYPHVQKWVREELNRELGAGQAPSLGDRTRLPYTDVHLHEA 360  
Db 301 FAGTMTVSTTVGYTLLLMKYPHVQKWVREELNRELGAGQAPSLGDRTRLPYTDVHLHEA 360  
QY 361 QRLALVPMGIPRTLMRTTRFRGYTLPGQTEVPFLGSLIHDNIFKHPEEFNPDRLDA 420  
Db 361 QRLALVPMGIPRTLMRTTRFRGYTLPGQTEVPFLGSLIHDNIFKHPEEFNPDRLDA 420  
QY 421 DGRFRKHEAFLPFLSKRVCLGEGAKAEFLFPTTILQAFSLBSPCPDPTLSLKPTVSG 480

Db 421 DGRFRKHEAFPLPSLGRKVCLEGLAKAEFLFPTTILQAFSLSPCPDPTLSLKPTVSG 480  
Qy 481 LFNIPPAFQLOVRPTDLHSTTTQTR 504  
Db 481 LFNIPPAFQLOVRPTDLHSTTTQTR 504

RESULT 9  
US-10-669-693-2  
; Sequence 2, Application US/10669693  
; Publication No. US20040043413A1  
; GENERAL INFORMATION:  
; APPLICANT: YAN, Chunhua et al  
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING  
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN  
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS,  
; TITLE OF INVENTION: AND USES THEREOF  
; FILE REFERENCE: CL000685CON  
; CURRENT APPLICATION NUMBER: US/10/669,693  
; CURRENT FILING DATE: 2003-09-25  
; PRIOR APPLICATION NUMBER: 09/748,127  
; PRIOR FILING DATE: 2000-12-27  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 504  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-669-693-2

Query Match 100.0%; Score 2615; DB 12; Length 504;  
Best Local Similarity 100.0%; Pred. No. 8.1e-253;  
Matches 504; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MEATGTWALLALALLLLTALSGTRARGHLPPTPLPLGNLLQLRPGALYSGLMRL 60  
Db 1 MEATGTWALLALALLLLTALSGTRARGHLPPTPLPLGNLLQLRPGALYSGLMRL 60

Qy 61 SKKYGPFVTIYLGWPWPVVVLVQEAIVREALGQQAEEFSGRGTVAMLEGTDFDGHGVFFSN 120  
Db 61 SKKYGPFVTIYLGWPWPVVVLVQEAIVREALGQQAEEFSGRGTVAMLEGTDFDGHGVFFSN 120

Qy 121 GERWRQLRKFTMLALRDLCMGKEGEBELIOAEARCLVETFOGTEGRPFDPSSLAAQATSN 180  
Db 121 GERWRQLRKFTMLALRDLCMGKEGEBELIOAEARCLVETFOGTEGRPFDPSSLAAQATSN 180

Qy 181 VVCSLLFGLRFSYEDKEFOAVVRAAGTLLGVSSQGGQTYEMFSWFLRPLPGPHKQLLHH 240  
Db 181 VVCSLLFGLRFSYEDKEFOAVVRAAGTLLGVSSQGGQTYEMFSWFLRPLPGPHKQLLHH 240

Qy 241 VSTLAFTVRQVOHQGNLDASGPARDLVDAFLKMAQEQNPGTEFTNKMLMTVIYLL 300  
Db 241 VSTLAFTVRQVOHQGNLDASGPARDLVDAFLKMAQEQNPGTEFTNKMLMTVIYLL 300

Qy 301 FAGTMTVSTTVGYTLLLMKYPHVQKWVREELNRELGAQAPSLGDRTRLPYTDVAVLHEA 360  
Db 301 FAGTMTVSTTVGYTLLLMKYPHVQKWVREELNRELGAQAPSLGDRTRLPYTDVAVLHEA 360

Qy 361 ORLLALVPMGIPRTLMTTRFRGYTLLPQGTVEVFPGLGSIILHDPNIFKHPEEFNPDRLDA 420  
Db 361 ORLLALVPMGIPRTLMTTRFRGYTLLPQGTVEVFPGLGSIILHDPNIFKHPEEFNPDRLDA 420

Qy 421 DGRFRKHEAFPLPSLGRKVCLEGLAKAEFLFPTTILQAFSLSPCPDPTLSLKPTVSG 480  
Db 421 DGRFRKHEAFPLPSLGRKVCLEGLAKAEFLFPTTILQAFSLSPCPDPTLSLKPTVSG 480

Qy 481 LFNIPPAFQLOVRPTDLHSTTTQTR 504  
Db 481 LFNIPPAFQLOVRPTDLHSTTTQTR 504

RESULT 10

US-10-140-024-64  
; Sequence 64, Application US/10140024  
; Publication No. US20040058424A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Bersini, Maureen  
; APPLICANT: DeForge, Laura  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Sherwood, Steven  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; TITLE OF INVENTION: ACIDS ENCODING THE SAME  
; FILE REFERENCE: P3330RIC69  
; CURRENT APPLICATION NUMBER: US/10/140,024  
; CURRENT FILING DATE: 2002-05-06  
; Prior Application removed - See Palm or File Wrapper  
; NUMBER OF SEQ ID NOS: 550  
; SEQ ID NO 64  
; LENGTH: 504  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-140-024-64

Query Match 100.0%; Score 2615; DB 12; Length 504;  
Best Local Similarity 100.0%; Pred. No. 8.1e-253;  
Matches 504; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MEATGTWALLALALLLLTALSGTRARGHLPPTPLPLGNLLQLRPGALYSGLMRL 60  
Db 1 MEATGTWALLALALLLLTALSGTRARGHLPPTPLPLGNLLQLRPGALYSGLMRL 60

Qy 61 SKKYGPFVTIYLGWPWPVVVLVQEAIVREALGQQAEEFSGRGTVAMLEGTDFDGHGVFFSN 120  
Db 61 SKKYGPFVTIYLGWPWPVVVLVQEAIVREALGQQAEEFSGRGTVAMLEGTDFDGHGVFFSN 120

Qy 121 GERWRQLRKFTMLALRDLCMGKEGEBELIOAEARCLVETFOGTEGRPFDPSSLAAQATSN 180  
Db 121 GERWRQLRKFTMLALRDLCMGKEGEBELIOAEARCLVETFOGTEGRPFDPSSLAAQATSN 180

Qy 181 VVCSLLFGLRFSYEDKEFOAVVRAAGTLLGVSSQGGQTYEMFSWFLRPLPGPHKQLLHH 240  
Db 181 VVCSLLFGLRFSYEDKEFOAVVRAAGTLLGVSSQGGQTYEMFSWFLRPLPGPHKQLLHH 240

Qy 241 VSTLAFTVRQVOHQGNLDASGPARDLVDAFLKMAQEQNPGTEFTNKMLMTVIYLL 300  
Db 241 VSTLAFTVRQVOHQGNLDASGPARDLVDAFLKMAQEQNPGTEFTNKMLMTVIYLL 300

Qy 301 FAGTMTVSTTVGYTLLLMKYPHVQKWVREELNRELGAQAPSLGDRTRLPYTDVAVLHEA 360  
Db 301 FAGTMTVSTTVGYTLLLMKYPHVQKWVREELNRELGAQAPSLGDRTRLPYTDVAVLHEA 360

Qy 361 ORLLALVPMGIPRTLMTTRFRGYTLLPQGTVEVFPGLGSIILHDPNIFKHPEEFNPDRLDA 420  
Db 361 ORLLALVPMGIPRTLMTTRFRGYTLLPQGTVEVFPGLGSIILHDPNIFKHPEEFNPDRLDA 420

Qy 421 DGRFRKHEAFPLPSLGRKVCLEGLAKAEFLFPTTILQAFSLSPCPDPTLSLKPTVSG 480  
Db 421 DGRFRKHEAFPLPSLGRKVCLEGLAKAEFLFPTTILQAFSLSPCPDPTLSLKPTVSG 480

Qy 481 LFNIPPAFQLOVRPTDLHSTTTQTR 504  
Db 481 LFNIPPAFQLOVRPTDLHSTTTQTR 504



RESULT 11  
US-10-140-808-64  
; Sequence 64, Application US/10140808  
; Publication No. US20030017563A1  
; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: DeForge, Laura  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Sherwood, Steven  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tamas, Daniel  
; APPLICANT: Watanabe, Colin K  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3330R1C182  
; CURRENT APPLICATION NUMBER: US/10/140,808  
; CURRENT FILING DATE: 2002-05-07  
; Prior Application removed - See File Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 550  
; SEQ ID NO 64  
; LENGTH: 504  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-140-808-64

Query Match 100.0%; Score 2615; DB 12; Length 504;  
Best Local Similarity 100.0%; Pred. No. 8.1e-253;  
Matches 504; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MEATGTWALLALALLLLTALSGTRARGHLPPTPLPLGNLLQLRPGALYSGLMRL 60  
DB 1 MEATGTWALLALALLLLTALSGTRARGHLPPTPLPLGNLLQLRPGALYSGLMRL 60  
QY 61 SKYGPVFTIYLGPNRPVVVLVGOEAVREALGQAEFFSGRTVAMLEGTFDGHGVFFSN 120  
DB 61 SKYGPVFTIYLGPNRPVVVLVGOEAVREALGQAEFFSGRTVAMLEGTFDGHGVFFSN 120  
QY 121 GERWQLRKFTMLALRDGLMGKREGBELIQAEARCLVETFGTEGPRPFDPSLLLAQATSN 180  
DB 121 GERWQLRKFTMLALRDGLMGKREGBELIQAEARCLVETFGTEGPRPFDPSLLLAQATSN 180  
QY 181 VVCSLLFGLRFSYEDKEFOAVVRAAGTLLGVSSQGGQTYEMFSWFLRPLPGPHKQLLHH 240  
DB 181 VVCSLLFGLRFSYEDKEFOAVVRAAGTLLGVSSQGGQTYEMFSWFLRPLPGPHKQLLHH 240  
QY 241 VSTLAFTVROVQOHOGLNDASGPARDLVDAFLKMAQEONPGTEFTNKNMLMTVIYLL 300  
DB 241 VSTLAFTVROVQOHOGLNDASGPARDLVDAFLKMAQEONPGTEFTNKNMLMTVIYLL 300  
QY 301 FAGTMTVSTVTGVTLLLMKYPHVQKWVREELNRELGAQAPSLGDRTRLPTDVLVHEA 360  
DB 301 FAGTMTVSTVTGVTLLLMKYPHVQKWVREELNRELGAQAPSLGDRTRLPTDVLVHEA 360  
QY 361 ORLLALVPNGIPRTLMTTRFRGYTLPGQTEVPPLLSILHDPNIFKHPEEENPDRFLDA 420  
DB 361 ORLLALVPNGIPRTLMTTRFRGYTLPGQTEVPPLLSILHDPNIFKHPEEENPDRFLDA 420  
QY 421 DGRFRKHEAFLPFLSGKRVCLGEGAKAELFLFTTILQAFSLSPCPDPTLSLKPTVSG 480  
DB 421 DGRFRKHEAFLPFLSGKRVCLGEGAKAELFLFTTILQAFSLSPCPDPTLSLKPTVSG 480

QY 481 LENIPPAFOQLVRPTDLHSTTQTR 504  
DB 481 LENIPPAFOQLVRPTDLHSTTQTR 504

## RESULT 12

US-10-152-405-64  
; Sequence 64, Application US/10152405  
; Publication No. US20030211571A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: DeForge, Laura  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Sherwood, Steven  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tamas, Daniel  
; APPLICANT: Watanabe, Colin K  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3330R1C383  
; CURRENT APPLICATION NUMBER: US/10/152,405  
; CURRENT FILING DATE: 2002-05-20  
; Prior Application removed - See File Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 550  
; SEQ ID NO 64  
; LENGTH: 504  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-152-405-64

Query Match 100.0%; Score 2615; DB 12; Length 504;  
Best Local Similarity 100.0%; Pred. No. 8.1e-253;  
Matches 504; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MEATGTWALLALALLLLTALSGTRARGHLPPTPLPLGNLLQLRPGALYSGLMRL 60  
DB 1 MEATGTWALLALALLLLTALSGTRARGHLPPTPLPLGNLLQLRPGALYSGLMRL 60  
QY 61 SKYGPVFTIYLGPNRPVVVLVGOEAVREALGQAEFFSGRTVAMLEGTFDGHGVFFSN 120  
DB 61 SKYGPVFTIYLGPNRPVVVLVGOEAVREALGQAEFFSGRTVAMLEGTFDGHGVFFSN 120  
QY 121 GERWQLRKFTMLALRDGLMGKREGBELIQAEARCLVETFGTEGPRPFDPSLLLAQATSN 180  
DB 121 GERWQLRKFTMLALRDGLMGKREGBELIQAEARCLVETFGTEGPRPFDPSLLLAQATSN 180  
QY 181 VVCSLLFGLRFSYEDKEFOAVVRAAGTLLGVSSQGGQTYEMFSWFLRPLPGPHKQLLHH 240  
DB 181 VVCSLLFGLRFSYEDKEFOAVVRAAGTLLGVSSQGGQTYEMFSWFLRPLPGPHKQLLHH 240  
QY 241 VSTLAFTVROVQOHOGLNDASGPARDLVDAFLKMAQEONPGTEFTNKNMLMTVIYLL 300  
DB 241 VSTLAFTVROVQOHOGLNDASGPARDLVDAFLKMAQEONPGTEFTNKNMLMTVIYLL 300  
QY 301 FAGTMTVSTVTGVTLLLMKYPHVQKWVREELNRELGAQAPSLGDRTRLPTDVLVHEA 360  
DB 301 FAGTMTVSTVTGVTLLLMKYPHVQKWVREELNRELGAQAPSLGDRTRLPTDVLVHEA 360  
QY 361 ORLLALVPNGIPRTLMTTRFRGYTLPGQTEVPPLLSILHDPNIFKHPEEENPDRFLDA 420  
DB 361 ORLLALVPNGIPRTLMTTRFRGYTLPGQTEVPPLLSILHDPNIFKHPEEENPDRFLDA 420  
QY 421 DGRFRKHEAFLPFLSGKRVCLGEGAKAELFLFTTILQAFSLSPCPDPTLSLKPTVSG 480

Db 61 SKYGPVFTIYLGWPRVWVLVGOEAVREALGQAEFFSGRGTVAMLEGTDFDGHGVFFSN 120  
QY 121 GERWRQKRFMTALRDGLMGKREGELIQAEARCLVETFOCTEGRPPDPDPSLLLAQAATSN 180  
Db 121 GERWRQKRFMTALRDGLMGKREGELIQAEARCLVETFOCTEGRPPDPDPSLLLAQAATSN 180  
QY 181 VVCSILLFGLRFSYEDKEFAVRAAGTTLGVSSQGGTYEMFSLPLPGPHKQLLHH 240  
Db 181 VVCSILLFGLRFSYEDKEFAVRAAGTTLGVSSQGGTYEMFSLPLPGPHKQLLHH 240  
QY 241 VSTLAFTVRQVQOHOQNDASGPARDLVDAFLKMAOEONPGTEFNKMLMTVIYLL 300  
Db 241 VSTLAFTVRQVQOHOQNDASGPARDLVDAFLKMAOEONPGTEFNKMLMTVIYLL 300  
QY 301 FAGTMTVSTTVGYTLLLMKYPHVQKVREELNRELGLGAGQAPSLGDRTELPTDVLHEA 360  
Db 301 FAGTMTVSTTVGYTLLLMKYPHVQKVREELNRELGLGAGQAPSLGDRTELPTDVLHEA 360  
QY 361 QRLALVPMGIPRTLMRTTRFRGYTLPGQTEVFVLLGSLILHDPNFKHPEENPDRFLDA 420  
Db 361 QRLALVPMGIPRTLMRTTRFRGYTLPGQTEVFVLLGSLILHDPNFKHPEENPDRFLDA 420  
QY 421 DGRFRKHEAFLPFSGLKRVCLGEGAKAELFLFFTTILQAFSLSPCPDPTLSLKPTVSG 480  
Db 421 DGRFRKHEAFLPFSGLKRVCLGEGAKAELFLFFTTILQAFSLSPCPDPTLSLKPTVSG 480  
QY 481 LFNIPPAFQLOVRPTDLHSTTQTR 504  
Db 481 LFNIPPAFQLOVRPTDLHSTTQTR 504

RESULT 14  
US-10-127-900A-64  
; Sequence 64, Application US/10127900A  
; Publication No. US20030203429A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: DeForge, Laura  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Sherwood, Steven  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3330R1C81  
; CURRENT APPLICATION NUMBER: US/10/127,900A  
; CURRENT FILING DATE: 2002-10-15  
; PRIOR APPLICATION NUMBER: 60/049911  
; PRIOR FILING DATE: 1997-06-18  
; PRIOR APPLICATION NUMBER: 60/056974  
; PRIOR FILING DATE: 1997-08-26  
; PRIOR APPLICATION NUMBER: 60/059113  
; PRIOR FILING DATE: 1997-09-17  
; PRIOR APPLICATION NUMBER: 60/059115  
; PRIOR FILING DATE: 1997-09-17  
; PRIOR APPLICATION NUMBER: 60/059117  
; PRIOR FILING DATE: 1997-09-17  
; PRIOR APPLICATION NUMBER: 60/059122  
; PRIOR FILING DATE: 1997-09-17  
; PRIOR APPLICATION NUMBER: 60/059184  
; PRIOR FILING DATE: 1997-09-17  
; PRIOR APPLICATION NUMBER: 60/059263

Db 421 DGRFRKHEAFLPFSGLKRVCLGEGAKAELFLFFTTILQAFSLSPCPDPTLSLKPTVSG 480  
QY 481 LFNIPPAFQLOVRPTDLHSTTQTR 504  
Db 481 LFNIPPAFQLOVRPTDLHSTTQTR 504  
RESULT 13  
US-10-127-852A-64  
; Sequence 64, Application US/10127852A  
; Publication No. US20030203428A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: DeForge, Laura  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Sherwood, Steven  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3330R1C88  
; CURRENT APPLICATION NUMBER: US/10/127,852A  
; CURRENT FILING DATE: 2002-10-15  
; PRIOR APPLICATION NUMBER: 60/049911  
; PRIOR FILING DATE: 1997-06-18  
; PRIOR APPLICATION NUMBER: 60/056974  
; PRIOR FILING DATE: 1997-08-26  
; PRIOR APPLICATION NUMBER: 60/059113  
; PRIOR FILING DATE: 1997-09-17  
; PRIOR APPLICATION NUMBER: 60/059115  
; PRIOR FILING DATE: 1997-09-17  
; PRIOR APPLICATION NUMBER: 60/059117  
; PRIOR FILING DATE: 1997-09-17  
; PRIOR APPLICATION NUMBER: 60/059122  
; PRIOR FILING DATE: 1997-09-17  
; PRIOR APPLICATION NUMBER: 60/059184  
; PRIOR FILING DATE: 1997-09-17  
; PRIOR APPLICATION NUMBER: 60/059263  
; PRIOR FILING DATE: 1997-09-18  
; PRIOR APPLICATION NUMBER: 60/059352  
; PRIOR FILING DATE: 1997-09-19  
; PRIOR APPLICATION NUMBER: 60/059588  
; PRIOR FILING DATE: 1997-09-19  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 550  
; SEQ ID NO 64  
; LENGTH: 504  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-127-852A-64

Query Match 100.0%; Score 2615; DB 12; Length 504;  
Best Local Similarity 100.0%; Pred. No. 8.1e-253;  
Matches 504; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MEATGTWALLALALLLLLTALSGTRARGHLPGPPTPLPLGNNLQLRPGALYSGLMRL 60  
Db 1 MEATGTWALLALALLLLLTALSGTRARGHLPGPPTPLPLGNNLQLRPGALYSGLMRL 60  
QY 61 SKYGPVFTIYLGWPRVWVLVGOEAVREALGQAEFFSGRGTVAMLEGTDFDGHGVFFSN 120

```
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059588
; PRIOR FILING DATE: 1997-09-19
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 64
; LENGTH: 504
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-127-900A-64

Query Match      100.0%; Score 2615; DB 12; Length 504;
Best Local Similarity 100.0%; Pred. No. 8.1e-253;
Matches 504; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEATGTWALLALALLLLTALSGTRARGHLPGPTPLPLGNNLQLRPGALYSGLMRL 60
DB 1 MEATGTWALLALALLLLTALSGTRARGHLPGPTPLPLGNNLQLRPGALYSGLMRL 60

QY 61 SKKYGPFVTYLGPRPVVVLVGOEAVREALGGOAEFSGRGTVAMLEGTFDGHGVFSN 120
DB 61 SKKYGPFVTYLGPRPVVVLVGOEAVREALGGOAEFSGRGTVAMLEGTFDGHGVFSN 120

QY 121 GERWRLRKFTMLALRDGLGKREGEELIOAEARCLVETFGTEGRPFDPSSLQAQTSN 180
DB 121 GERWRLRKFTMLALRDGLGKREGEELIOAEARCLVETFGTEGRPFDPSSLQAQTSN 180

QY 181 VCSLLFGLRFSYEDKEFQAVVRAAGTLLGVSSQGGQTYEMFSWFLRPLPGPHKQLLHH 240
DB 181 VCSLLFGLRFSYEDKEFQAVVRAAGTLLGVSSQGGQTYEMFSWFLRPLPGPHKQLLHH 240

QY 241 VSTLAFTVRQVQOQHONLDASGPARDLVDAFLKMAQEEQNPGTEFTNKNMLMTVIYLL 300
DB 241 VSTLAFTVRQVQOQHONLDASGPARDLVDAFLKMAQEEQNPGTEFTNKNMLMTVIYLL 300

QY 301 FAGTMTVSTTVGYTLLLMKYPHVQKWVREELNRELGAGAPSLGDRTRLPYTDVILHEA 360
DB 301 FAGTMTVSTTVGYTLLLMKYPHVQKWVREELNRELGAGAPSLGDRTRLPYTDVILHEA 360

QY 361 QRLALVPMGIPRTLMTTRFRGYTLPOGTEVPFLGSLIHDNPIFKHPEEFNPDRLDA 420
DB 361 QRLALVPMGIPRTLMTTRFRGYTLPOGTEVPFLGSLIHDNPIFKHPEEFNPDRLDA 420

QY 421 DGRFRKHEAFLPSLGKRVCLGELAKAELFLFTTILQAFSLESPPDPTLSLKPTVSG 480
DB 421 DGRFRKHEAFLPSLGKRVCLGELAKAELFLFTTILQAFSLESPPDPTLSLKPTVSG 480

QY 481 LFNIPPAFQVLRPTDLHSTTQTR 504
DB 481 LFNIPPAFQVLRPTDLHSTTQTR 504
```

## RESULT 15

US-10-128-685A-64

; Sequence 64, Application US/10128685A

; Publication No. US20030203430A1

## GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Beresini, Maureen

; APPLICANT: DeForge, Laura

; APPLICANT: Deenoyers, Luc

; APPLICANT: Filvaroff, Ellen

; APPLICANT: Gao, Wei-Qiang

; APPLICANT: Gerritsen, Mary E.

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Sherwood, Steven

; APPLICANT: Smith, Victoria

; APPLICANT: Stewart, Timothy A.

; APPLICANT: Tuma, Daniel

```
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C116
; CURRENT APPLICATION NUMBER: US/10/128,685A
; CURRENT FILING DATE: 2002-04-23
; PRIOR APPLICATION NUMBER: 60/049911
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: 60/056974
; PRIOR FILING DATE: 1997-08-26
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059115
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059117
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059122
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059184
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059588
; PRIOR FILING DATE: 1997-09-19
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 64
; LENGTH: 504
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-128-685A-64
```

Query Match 100.0%; Score 2615; DB 12; Length 504;

Best Local Similarity 100.0%; Pred. No. 8.1e-253;

Matches 504; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEATGTWALLALALLLLTALSGTRARGHLPGPTPLPLGNNLQLRPGALYSGLMRL 60

DB 1 MEATGTWALLALALLLLTALSGTRARGHLPGPTPLPLGNNLQLRPGALYSGLMRL 60

QY 61 SKKYGPFVTYLGPRPVVVLVGOEAVREALGGOAEFSGRGTVAMLEGTFDGHGVFSN 120

DB 61 SKKYGPFVTYLGPRPVVVLVGOEAVREALGGOAEFSGRGTVAMLEGTFDGHGVFSN 120

QY 121 GERWRLRKFTMLALRDGLGKREGEELIOAEARCLVETFGTEGRPFDPSSLQAQTSN 180

DB 121 GERWRLRKFTMLALRDGLGKREGEELIOAEARCLVETFGTEGRPFDPSSLQAQTSN 180

QY 181 VCSLLFGLRFSYEDKEFQAVVRAAGTLLGVSSQGGQTYEMFSWFLRPLPGPHKQLLHH 240

DB 181 VCSLLFGLRFSYEDKEFQAVVRAAGTLLGVSSQGGQTYEMFSWFLRPLPGPHKQLLHH 240

QY 241 VSTLAFTVRQVQOQHONLDASGPARDLVDAFLKMAQEEQNPGTEFTNKNMLMTVIYLL 300

DB 241 VSTLAFTVRQVQOQHONLDASGPARDLVDAFLKMAQEEQNPGTEFTNKNMLMTVIYLL 300

QY 301 FAGTMTVSTTVGYTLLLMKYPHVQKWVREELNRELGAGAPSLGDRTRLPYTDVILHEA 360

DB 301 FAGTMTVSTTVGYTLLLMKYPHVQKWVREELNRELGAGAPSLGDRTRLPYTDVILHEA 360

QY 361 QRLALVPMGIPRTLMTTRFRGYTLPOGTEVPFLGSLIHDNPIFKHPEEFNPDRLDA 420

DB 361 QRLALVPMGIPRTLMTTRFRGYTLPOGTEVPFLGSLIHDNPIFKHPEEFNPDRLDA 420

QY 421 DGRFRKHEAFLPSLGKRVCLGELAKAELFLFTTILQAFSLESPPDPTLSLKPTVSG 480

DB 421 DGRFRKHEAFLPSLGKRVCLGELAKAELFLFTTILQAFSLESPPDPTLSLKPTVSG 480

QY 481 LFNIPPAFQVLRPTDLHSTTQTR 504

Db 481 LFNIPPAFQLOVRPTDLHSTTQTR 504

Search completed: September 15, 2004, 09:33:07  
Job time : 130 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2004, 09:18:40 ; Search time 33 Seconds

(without alignments)  
788.469 Million cell updates/sec

Title: US-10-669-693-2

Perfect score: 2615

Sequence: 1 MEATGTWALLALALLLT.....PPAFQLQVRPTDLHSTTQTR 504

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%

Listing first 45 summaries

Database :

Issued Patents AA.\*  
1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pap.\*  
2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pap.\*  
3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pap.\*  
4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pap.\*  
5: /cgn2\_6/ptodata/2/iaa/PTCUS\_COMB.pap.\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1231.5	47.1	494	US-09-126-420A-26	Sequence 26, Appl
2	1092.5	41.8	490	US-08-201-118-13	Sequence 13, Appl
3	1092.5	41.8	490	US-08-238-821B-13	Sequence 13, Appl
4	1092.5	41.8	490	PCT-US95-05744-13	Sequence 13, Appl
5	1087.5	41.6	490	US-08-201-118-3	Sequence 3, Appl
6	1087.5	41.6	490	US-08-238-821B-3	Sequence 3, Appl
7	1087.5	41.6	490	PCT-US95-05744-3	Sequence 3, Appl
8	1085.5	41.5	490	US-08-201-118-9	Sequence 9, Appl
9	1085.5	41.5	490	US-08-238-821B-9	Sequence 9, Appl
10	1085.5	41.5	490	PCT-US95-05744-9	Sequence 9, Appl
11	1084.5	41.5	490	US-08-201-118-1	Sequence 1, Appl
12	1084.5	41.5	490	US-08-238-821B-1	Sequence 1, Appl
13	1084.5	41.5	490	PCT-US95-05744-1	Sequence 1, Appl
14	1073.5	41.1	490	US-08-201-118-5	Sequence 5, Appl
15	1073.5	41.1	490	US-08-238-821B-5	Sequence 5, Appl
16	1073.5	41.1	490	PCT-US95-05744-5	Sequence 5, Appl
17	1070.5	40.9	500	US-08-314-601-2	Sequence 2, Appl
18	1070.5	40.9	500	PCT-US95-13051-2	Sequence 2, Appl
19	1067.5	40.8	490	US-08-201-118-11	Sequence 11, Appl
20	1067.5	40.8	490	US-08-238-821B-11	Sequence 11, Appl
21	1067.5	40.8	490	PCT-US95-05744-11	Sequence 11, Appl
22	1050.5	40.2	490	US-08-201-118-7	Sequence 7, Appl
23	1050.5	40.2	490	US-08-238-821B-7	Sequence 7, Appl
24	1050.5	40.2	490	PCT-US95-05744-7	Sequence 7, Appl
25	792	30.3	504	US-09-126-420A-23	Sequence 23, Appl
26	679	26.0	502	US-09-304-615-69	Sequence 69, Appl
27	562.5	21.5	512	US-08-194-981E-5	Sequence 5, Appl

28	544.5	20.8	524	4	US-09-126-420A-24	Sequence 24, Appl
29	502	19.2	504	4	US-09-976-594-642	Sequence 642, App
30	439	16.8	508	4	US-09-126-420A-25	Sequence 25, Appl
31	421	16.1	501	3	US-09-158-767-15	Sequence 15, Appl
32	421	16.1	501	3	US-09-158-767-16	Sequence 16, Appl
33	421	16.1	501	3	US-09-158-767-17	Sequence 17, Appl
34	421	16.1	501	3	US-09-158-767-18	Sequence 18, Appl
35	417	15.9	504	4	US-09-499-302A-9	Sequence 9, Appl
36	405.5	15.5	506	4	US-09-672-785-2	Sequence 2, Appl
37	401.5	15.4	490	4	US-09-126-420A-16	Sequence 16, Appl
38	386.5	14.8	476	1	US-08-313-075A-30	Sequence 30, Appl
39	378	14.5	471	4	US-09-126-420A-20	Sequence 20, Appl
40	378	14.5	495	1	US-08-532-065B-2	Sequence 2, Appl
41	377	14.4	506	4	US-09-672-785-8	Sequence 8, Appl
42	374.5	14.3	505	4	US-09-627-216A-12	Sequence 12, Appl
43	374.5	14.3	505	4	US-09-765-873A-12	Sequence 12, Appl
44	374	14.3	531	4	US-09-380-420C-2	Sequence 2, Appl
45	374	14.3	531	4	US-09-899-642A-2	Sequence 2, Appl

## ALIGNMENTS

RESULT 1  
US-09-126-420A-26  
; Sequence 26, Application US/09126420A  
; Patent No. 6376753  
; GENERAL INFORMATION:  
; APPLICANT: BATARD, YANNICK  
; APPLICANT: ROBINEAU, TIBURCE  
; APPLICANT: DUREST, FRANCIS  
; APPLICANT: WERCK-REICHART, DANIELE  
; APPLICANT: DIDIERJEAN, LUC  
; TITLE OF INVENTION: PURIFIED CYTOCHROME P450 CYP76B1 FROM HELIANTHUS  
; TITLE OF INVENTION: TUBEROSUS AND ITS APPLICATIONS AS BIOCATALYST IN  
; TITLE OF INVENTION: PARTICULAR FOR THE DEGRADATION OF ENVIRONMENTAL  
; TITLE OF INVENTION: POLLUTANTS AND FOR ALTERING THE RESISTANCE OF PLANTS S  
; FILE REFERENCE: 03715.0032  
; CURRENT APPLICATION NUMBER: US/09/126.420A  
; CURRENT FILING DATE: 1998-07-30  
; PRIOR APPLICATION NUMBER: 60/054.351  
; PRIOR FILING DATE: 1997-07-31  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 26  
; LENGTH: 494  
; TYPE: PRT  
; ORGANISM: Rattus norvegicus  
US-09-126-420A-26

Query Match  
Best Local Similarity 47.1%; Score 1231.5; DB 4; Length 494;  
Matches 242; Conservative 93; Mismatches 150; Indels 9; Gaps 4;  
QY 1 MEATGTWALLALALLLTALSGT-----RARGHLPPGPTPLPLGNLQLRPGALYS 56  
Db 1 MTASG---LLVASVAFLSVLMSVWKQKLSKLPPTPLPFGNYQLNTERKMYSS 57  
QY 57 LMLSKYGVFTYILGWPVPVVLVGOEAVRGLGQAEFGSGRGTVAMLEGTFFDHGV 116  
Db 58 LMAISORYGVFTIHLGP-RRVVLCGQAEKALVDQAEFGSGRGQATFDFLFGYGV 116  
QY 117 FFSNGERWQLRKTFTMLALDLGMKREGEELIQAERCLVETFOGTEGRPFPPSLLAQ 176  
Db 117 AFSSGERAQLRFSFATLRDFGVKRGKEERLQEEAGFLIESFRKNGALIDFTYLSR 176  
QY 177 ATSNVCSLLFLGRFSYEDKFEQAVVRAAGTLLGVSSQGGQTYEMFSWFLRPLGPHKQ 236  
Db 177 TVSNVISSVFGDRFDYEDKEFLSLRLMWLGSFQFATSTGTQLYEMFSSWMKHLPGPQQ 236  
QY 237 LLHHVSTLAFTVROVQOHOGLDASGPARDLVDAFLKQAOEEQNPGTFTNNMLMTV 296  
Db 237 AFKEQGLEDFITTKVQEQNRTLDNPSR-RDFIDSLFIRMLEKKNPNTFYMKNLVLT 295

```
QY 297 IYLLFAGTMTYSTTGYTLLLLMKYPHVQKWRBELNRELGAGQAPSLGDRTRLPYTDAY 356
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 296 LNLFPAGTETVSTLRYGFLLLMRHPDIEAKVHEIDRVIGRNRQAKYEDRMKMPYTEAV 355
QY 357 LHEAQRLALVPMGIPRLMTTRRGTYTLPOGTEVFFLLGSLILHDPNIEKHPEEFNPDR 416
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 356 IHEIQRFADMI PMGLARRVTKDFREFLLPKGTVEFFMLGSLVLPKFFSNPNDFNPKH 415
QY 417 FLDADGRFRKEAFUPEFSLGKVCVLAGEGLAKAELELFFTTILQAFSLSPCPPDPTLSLKP 476
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 416 FLDDKGQPKKSDAFVFFSIGKRYCGEGGLARMELFLFLTNIMQNECFKSPQAPQIDVSP 475
QY 477 TVSGLFNIPPAFQL 490
:|||||:
Db 476 RLVGATIPENYTM 489

RESULT 2
US-08-201-118-13
; Sequence 13, Application US/08201118
; Patent No. 5786191
; GENERAL INFORMATION:
; APPLICANT: GOLDSTEIN, Joyce A.
; APPLICANT: ROMKES-SPARKS, Marjorie
; TITLE OF INVENTION: CLONING AND EXPRESSION OF COMPLEMENTARY
; TITLE OF INVENTION: DNAS FOR MULTIPLE MEMBERS OF THE HUMAN CYTOCHROME P450 2C
; TITLE OF INVENTION: SUBFAMILY
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Khourie and Crew
; STREET: 379 Lytton Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: US
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/201,118
; FILING DATE: 22-FEB-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/864,962
; FILING DATE: 09-APR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Liebeschuetz, Joe
; REGISTRATION NUMBER: 37,505
; REFERENCE/DOCKET NUMBER: 15280-192-1
; TELEPHONE: (415) 326-2400
; TELEFAX: (415) 326-2422
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 490 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; US-08-201-118-13
; Query Match 41.8%; Score 1092.5; DB 1; Length 490;
; Best Local Similarity 45.0%; Pred. No. 2.6e-107;
; Matches 217; Conservative 92; Mismatches 170; Indels 3; Gaps 3;
QY 9 LLLALALLLLTALSCTRAGHLPPOGTPPLPLGNLLQLRPPGALYGLMRLSKYGVF 68
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 7 LVLCLSCLLLSLRQSS-GRGKLPFGPTPLPXIGNILQIDKDISKSLTNKXKVGPF 65
```





Db 304 TTLVALLLLKHPVTAQVQEEIERTVIGNRKSPQMQRSHMPYTDVAVVHVQRYLDLLP 363  
Qy 369 MGIPRTLMRTFRGYTLPOGTEVFFPLGLSLHDPNIFKHPBEFNDRFLDADGRFKHE 428  
Db 364 TSLPHAVTCDIKFRNYLIPKGTTLILISLTVLHDKFEPNEMFDPHFLDEGGNFKSK 423  
Qy 429 AFLPESLGRKVCLEGLAKAELEFFFTTILQAFSLSPCPDPTLSLKPTVSGLFNIPAF 488  
Db 424 YFMFSAKRICVGEALAGMELFLFLTSILQNFNLKSLVDPKDLDTTPVAVGFASVPPEY 483  
Qy 489 QL 490  
Db 484 QL 485  
RESULT 6  
US-08-238-821B-3  
; Sequence 3, Application US/08238821B  
; Patent No. 5912120  
; GENERAL INFORMATION:  
; APPLICANT: GOLDSTEIN, Joyce A.  
; APPLICANT: ROMKES-SPARKS, Marjorie  
; TITLE OF INVENTION: CLONING, EXPRESSION AND DIAGNOSIS OF HUMAN  
; TITLE OF INVENTION: CYTOCHROME P450 2C19: THE PRINCIPAL DETERMINANT OF S-  
; TITLE OF INVENTION: MEPHENYTOIN METABOLISM  
; NUMBER OF SEQUENCES: 61  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, 8th Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: US  
; ZIP: 94111  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; FILING DATE: 06-MAY-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION NUMBER: US/08/238,821B  
; APPLICATION NUMBER: US 08/201,118  
; FILING DATE: 22-FEB-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/864,962  
; FILING DATE: 09-APR-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Liebeschuetz, Joe  
; REGISTRATION NUMBER: 37,505  
; REFERENCE/DOCKET NUMBER: 15280-192110US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (650) 326-2400  
; TELEFAX: (650) 326-2422  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 490 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-238-821B-3

Query Match 41.6%; Score 1087.5; DB 1; Length 490;  
Best Local Similarity 44.8%; Pred. No. 8.9e-107;  
Matches 216; Conservative 99; Mismatches 164; Indels 3; Gaps 3;  
Qy 9 LLLALALLLTLALSGTRAGHLPQGTPLPLGLNLLQLRPGALYSGLMRLSKYGPVF 68  
Db 7 LVLCSLCLLLSLWRQSS-GRGKLPQGTPLPVIGNILQIGIKDLSKLTNLSKVYGPVF 65  
Qy 69 TIYGLPWRPVVLVQGVAVREALGGAQAEPSRGRTVMLEGTGCHGVFSPSNERWQLR 128  
Db 66 TLYFG-LKPIVLVHGYEAVKEALIDGEEFSGRGIPLAERANRGFIVSNGKKWKEIR 124  
Qy 129 KFTMLALRDLGMGRGEEELIQAEARCLVETFGTEGRPDPDPLLLAQATSNVVCSLFFG 188  
Db 125 RFSMLTLRNGMGKRSIEDRVQEARCLVLEAKTKASPCDPTFILGCAPCNVICSIIFH 184  
Qy 189 LRFSEYDEKFAVVAAGTLLGVSSQGGQTYEMFSLRPLGPHKQLLHVSTLAFT 248  
Db 185 KRFYKQDQQLNLMKLNENIKILSPWICNNFSPDIIDYFPGTHNKLKLVAFMKSYI 244  
Qy 249 VRQVQHQGNLDASGPARDLVDAFLKWAQEQNPCTEFTKNMLMTVILLFAGTWTVS 308  
Db 245 LEKVKEQESMDMNNP-QDFIDCFMLKMKVEKHNQPSFTIESLENTAVDLFGAGTETS 303  
Qy 309 TTGVYTLTLLMKYPHVQKWRREELNRELGAQAPSLGDRTRLPYTDVLAHQELALVP 369

Qy	69	TIYGPWPRVVVLVQGEAVREALGGQABEEFSGRGTVAMLEGTGDHGVFFSNGERWROLR	128
Db	66	TLTYG-LKPIVVLHGVEAKALIDLGEESGRGIFPLAERANRGFIVFSNGKKWKEIR	124
Qy	129	KFTMLAURDLGMRKEGELIQAEARCIVETFOGTGEPFDSLLLAQNTSNVVCSSLPG	188
Db	125	RFSMTLNFNGKRSIDRVOEARCIVELBKTKASPCDFTFLGCAPCNVICSIIFH	184
Qy	189	LRFSYEDKEFOAVRAAGTTLGVSSQGGQTYEMESWFLRPLPGPHKLLHHVSTLAAFT	248
Db	185	KRFDYKDOQFLNMLEKNENIKILSSPMWIOICNNFSPIDIYPGTHNKLKNVAPMKSVI	244
Qy	249	VRVQOQHGNLDASPARDLVDALFKAAQEBQNGTFTNNQMLMTVIYLLFAGTMVTS	308
Db	245	LEKVKEHQSMDMNP-QDFIDCFMLKMKKEKHNPSEFTIESLENTAVDLFGAGTETTS	303
Qy	309	TTVGYTLLLLMKYPHVKWVREELNRELGAGQASLGDRTRLPYTDVILHEAQRLLALVP	368
Db	304	TLRLYALLLLLKHPEVTAKVOEEETVIGRNRSPQWRSHMPYTDVAVHEQVRYLDLLP	363
Qy	369	MGIPRTRLMRTRFRGYTLPGQTEVPFLLGSILHDPNIFKHPEEFNDRFLDADGRFKHE	428
Db	364	TSLPHAVTCDIKFRNYLIPKGTITLISLTVLHDNKEFPNPEMPDPHFHLDGEGNFKSK	423
Qy	429	AFLPFSLGKVCGLGEGLAKALELFFFTTLQAFSLSEPCPDYTLKKTVSGLFNIPAF	488
Db	424	YFNPFSNGKRCVGEALAGMELFLUTLSILQNFNLKSLVDPKNLDTTTPVNVNGFASVPFFY	483
Qy	489	QL 490	
Db	484	QL 485	

RESULT 7  
PCT-US95-05744-3  
; Sequence 3, Application PC/TUS9505744  
; GENERAL INFORMATION:  
; APPLICANT: GOLDSTEIN, Joyce A.  
; APPLICANT: ROMKES-SPARKS, Marjorie  
; APPLICANT: DE MORALS, Sonia M.F.  
; TITLE OF INVENTION: CLONING, EXPRESSION AND DIAGNOSIS OF HUMAN  
; TITLE OF INVENTION: CYTOCHROME P450 2C19: THE PRINCIPAL DETERMINANT  
; TITLE OF INVENTION: OF S-MEPHYNYTOIN METABOLISM  
; NUMBER OF SEQUENCES: 61  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend Khourie and Crew  
; STREET: 379 Lytton Avenue  
; CITY: Palo Alto  
; STATE: California  
; COUNTRY: US  
; ZIP: 94301  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/05744  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/238,821  
; FILING DATE: 06-MAY-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/201,118  
; FILING DATE: 22-FEB-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/864,962  
; FILING DATE: 09-APR-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Dow, Karen B.  
; REGISTRATION NUMBER: 29,584

```

: REFERENCE/DOCKET NUMBER: 15280-192-1-1-1
:
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (415) 326-2400
: TELEFAX: (415) 326-2422
: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 490 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: PCT-US95-05744-3

```

[illegible]

RESULT 8  
US-08-201-118-9  
; Sequence 9, Application US/08201118  
; Patent No. 5786191  
; GENERAL INFORMATION:  
; APPLICANT: GOLDSTEIN, Joyce A.  
; APPLICANT: ROMKES-SPARKS, Marjorie  
; TITLE OF INVENTION: CLONING AND EXPRESSION OF COMPLEMENTARY  
; TITLE OF INVENTION: DNAS FOR MULTIPLE MEMBERS OF THE HUMAN CYTOCHROME P450 2C  
; TITLE OF INVENTION: SUBFAMILY  
; NUMBER OF SEQUENCES: 44  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend Kourie and Crew  
; STREET: 379 Lytton Avenue  
; CITY: Palo Alto  
; STATE: California  
; COUNTRY: US  
; ZIP: 94301  
; COMPUTER READABLE FORM:

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/201,118
; FILING DATE: 22-FEB-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/864,962
; FILING DATE: 09-APR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Liebeschuetz, Joe
; REGISTRATION NUMBER: 37,505
; REFERENCE/DOCKET NUMBER: 15280-192-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 326-2400
; TELEFAX: (415) 326-2422
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 490 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-201-118-9

Query Match 41.5%; Score 1085.5; DB 1; Length 490;
Best Local Similarity 44.6%; Pred. No. 1.5e-106;
Matches 215; Conservative 100; Mismatches 164; Indels 3; Gaps 3;

QY 9 LLLALLLLLLTALSGRARHLPPTPLPLGNLLQLRPGALYSGLMRLSKKYGVPV 68
DB 7 LVLCSCLLLSLWRQSS-GRGKLPPGPTPLPVIGNILQIGKIDSKSLTNLSKVYGPV 65
QY 69 TTYLGPWRPVVVLVGOEAREALGQAEFSGRGTVAMLEGTFDGHGVFSGNWRQLR 128
DB 66 TLYFG-LKPIVVLHGVEAVKEALIDLGEFSGRGIFPLAERANRGFIVFSGKWKKEIR 124
QY 129 KFTMLALRDLGMKRGEBELIQAEARCLVETFTQTEGRPFDPSSLQAQATSNVVCLLFG 188
DB 125 RFSMTLRNFGMGKRSIEDRVQEARCLVEELRKTASPCDPTFILGCPCNVICSIIHF 184
QY 189 LRFSEYDEKFOAVRAAGTLLGVSSGGQTYEMFSWFLRPLGPHKQLLHVSTLAAFT 248
DB 185 KRFDYKQQOQFLNLMKLNENIKILSSPWIOICNNFSPIIDYPPGTHKLLKNVAFMKSYI 244
QY 249 VRQVOHQGNLDASGPARDLVDAFLKMAQEEONPGTEFTNKMLMTVIYLLFAGTMTVS 308
DB 245 LEKVKEHQESMDMNNP-QDFIDCFLMKMEKEKHQSEFTIESLENTAVDLFGAGTETTS 303
QY 309 TTVGYTLLLLMKYPHYQKWRBELNRELGAQAPSLGDRTRLPYTDVAVLHEAQRLLALVP 368
DB 304 TLRVALLLLKHPEVTAKVQEEIERVIGNRSPCMQDRSHMPYTDVAVHEVQRYIDL 363
QY 369 MGIPRLMTTRFRGYTLQCTGVFPLGSLHDPMIEKHPENFDRFLDADGRFRKHE 428
DB 364 TSLPHAVTCDIKFRNLYLIPKGTILLSLTSLVLDHNEFPENPFDDPHHLEDCGNFKSK 423
QY 429 AFLPFLSGKRVCLGEGLAKAEFLFTTILQAFSLSPCPDPTLSLKTPTVSGLFNIPPAF 488
DB 424 YFMPFSAGKRCVGEALAGNELFLTSLILQNFNLKSLVDPKNLDITPPVNGFASVPFFY 483
QY 489 QL 490
DB 484 QL 485

RESULT 9
US-08-238-821B-9
; Sequence 9, Application US/08238821B
; Patent No. 5912120
; GENERAL INFORMATION:
; APPLICANT: GOLDSTEIN, Joyce A.
; APPLICANT: ROMKES-SPARKS, Marjorie
; APPLICANT: DE MORALS, Sonia M.F.
; TITLE OF INVENTION: CLONING, EXPRESSION AND DIAGNOSIS OF HUMAN
; TITLE OF INVENTION: CYTOCHROME P450 2C19: THE PRINCIPAL DETERMINANT OF S-
; TITLE OF INVENTION: MEPHENYTOIN METABOLISM
; NUMBER OF SEQUENCES: 61
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: US
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/238,821B
; FILING DATE: 06-MAY-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/201,118
; FILING DATE: 22-FEB-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/864,962
; FILING DATE: 09-APR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Liebeschuetz, Joe
; REGISTRATION NUMBER: 37,505
; REFERENCE/DOCKET NUMBER: 15280-192110US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 326-2400
; TELEFAX: (650) 326-2422
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 490 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-238-821B-9

Query Match 41.5%; Score 1085.5; DB 2; Length 490;
Best Local Similarity 44.6%; Pred. No. 1.5e-106;
Matches 215; Conservative 100; Mismatches 164; Indels 3; Gaps 3;

QY 9 LLLALLLLLLTALSGRARHLPPTPLPLGNLLQLRPGALYSGLMRLSKKYGVPV 68
DB 7 LVLCSCLLLSLWRQSS-GRGKLPPGPTPLPVIGNILQIGKIDSKSLTNLSKVYGPV 65
QY 69 TTYLGPWRPVVVLVGOEAREALGQAEFSGRGTVAMLEGTFDGHGVFSGNWRQLR 128
DB 66 TLYFG-LKPIVVLHGVEAVKEALIDLGEFSGRGIFPLAERANRGFIVFSGKWKKEIR 124
QY 129 KFTMLALRDLGMKRGEBELIQAEARCLVETFTQTEGRPFDPSSLQAQATSNVVCLLFG 188
DB 125 RFSMTLRNFGMGKRSIEDRVQEARCLVEELRKTASPCDPTFILGCPCNVICSIIHF 184
QY 189 LRFSEYDEKFOAVRAAGTLLGVSSGGQTYEMFSWFLRPLGPHKQLLHVSTLAAFT 248
DB 185 KRFDYKQQOQFLNLMKLNENIKILSSPWIOICNNFSPIIDYPPGTHKLLKNVAFMKSYI 244
QY 249 VRQVOHQGNLDASGPARDLVDAFLKMAQEEONPGTEFTNKMLMTVIYLLFAGTMTVS 308
DB 245 LEKVKEHQESMDMNNP-QDFIDCFLMKMEKEKHQSEFTIESLENTAVDLFGAGTETTS 303
QY 309 TTVGYTLLLLMKYPHYQKWRBELNRELGAQAPSLGDRTRLPYTDVAVLHEAQRLLALVP 368
DB 304 TLRVALLLLKHPEVTAKVQEEIERVIGNRSPCMQDRSHMPYTDVAVHEVQRYIDL 363
```

Qy	369	MGIPTLMTTRFRGYTLPOGTEVFPLIGSLIHDPNIFKHPEEFNPDRFLDADGRPKHE	428
	:	: : : : : :	:
D <sub>b</sub>	364	TSLPHAVTCDIKFRNYLPKPTTLLISITSVLHDNKEFPNPFDPHHFDEGGNFKKSK	423
	:	: : : : : :	:
Qy	429	AFLPESLGKRVCLGEGLAKAELFFFTTLLQAFLSESPCPDTLSLKPTVSGLFNIPPAF	488
	:	: : : : : :	:
D <sub>b</sub>	424	YFMPEPSAGKRVCVEALAGMELFLFLTSLQNFNLSLVDPKNLDITTPVVGFAVSVPPEY	483
	:	: : : : : :	:
Qy	489	QL 490	
	:	: : : : : :	:
D <sub>b</sub>	484	OL 485	
	:	: : : : : :	:

RESULT 10  
PCT-US95-05744-9  
; Sequence 9, Application PC/TUS9505744  
; GENERAL INFORMATION:  
; APPLICANT: GOLDSTEIN, Joyce A.  
; APPLICANT: ROMKES-SPARKS, Marjorie  
; APPLICANT: DE MORAIS, Sonia M.F.  
; TITLE OF INVENTION: CLONING, EXPRESSION AND DIAGNOSIS OF HUMAN  
; TITLE OF INVENTION: CYTOCHROME P450 2C19: THE PRINCIPAL DETERMINANT  
; TITLE OF INVENTION: OF S-MEPHNYNTIN METABOLISM  
; NUMBER OF SEQUENCES: 61  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend Khourie and Crew  
; STREET: 379 Lytton Avenue  
; CITY: Palo Alto  
; STATE: California  
; COUNTRY: US  
; ZIP: 94301  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/05744  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/238,821  
; FILING DATE: 06-MAY-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/201,118  
; FILING DATE: 22-FEB-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/864,962  
; FILING DATE: 09-APR-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Dow, Karen B.  
; REGISTRATION NUMBER: 29,684  
; REFERENCE/DOCKET NUMBER: 15280-192-1-1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 326-2400  
; TELEFAX: (415) 326-2422  
; INFORMATION FOR SEQ ID NO: 9:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 490 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
PCT-US95-05744-9

```

QY 69 TTYLGPWRPVVVLVGQEAAREALGCGQAEEFSGRGIVAMLEGTFDGHGVFFSNGERWQLR 128
DB 66 TLYFG-LKPIVVVLHGYEAVKEALIDLGEFSGRGIFPLAERANRGFIVPSNGKKWKEIR 124
QY 129 KFTMLAIRDLMGKREGBELIOABEARCLVETFOGTGREGFPDPSLLLAQAATSNVVCSSLFG 188
DB 125 RFSLMTLRFNGMGRKSIEDRVQEEARCLVEELRKTAKSPCDPTFILGCAPCNVICSIIFH 184
QY 189 LRFSEYDEKEFOAVVRAAGCTLLGVSSQGGOTVEMFSEWFLRPLPGPHKOLLHHVSTLAAFT 248
DB 185 KRFDYKQOQLFLNMEKLNENIKILSSPWIQICNFFSPIIDYFGTHNKLKNVAFMKSYI 244
QY 249 VRQVOHQGNLDAGSPARDLVDAFLKMAQEONFCTEFTNKNMLMTVIYLLFAGTMTVS 308
DB 245 LEKVKEHQESMDMNP-QDFIDCFMLKMEKEKHNPSEFTIESLENTAVDLFGAGTETTS 303
QY 309 TTVGYTLLLMKYPHVQKVVREELNRELGAGAPSLGDRTRLPYTDVHEAQRLLALVP 368
DB 304 TTYRYALLLKHPEVAKVQEEIERVGRNRCPCMQDRSHMPYTDVAVVHEVQRYIDLPL 363
QY 369 MGIPRTLMRTTRTFRGYTLPOQTEVFPFLGSLIHDNPIFKHPEENPDRFLDADGRERKHE 428
DB 364 TSLPHAVTCDKFENYILPKGTILISLTVLHDNKEFPNPEMFPDHHFDEGGNFKKSK 423
QY 429 AFIPFSLGKRVCLGEGAKALBELFUFFTTILOAFSLESPPDPTLSLKPTVSGLFNIPPAF 488
DB 424 YMFPSAGKRICVGEALAGMELFLTSLILQNFNLKSLVDPKNLDTPVWVNGFASVPPFY 483
QY 489 QL 490
DB 484 QL 485

RESULT 11
US-08-201-118-1
: Sequence 1, Application US/08201118
: Patent No. 5786191
: GENERAL INFORMATION:
: APPLICANT: GOLDSTEIN, Joyce A.
: APPLICANT: ROMKES-SPARKS, Marjorie
: TITLE OF INVENTION: CLONING AND EXPRESSION OF COMPLEMENTARY
: TITLE OF INVENTION: DNAs FOR MULTIPLE MEMBERS OF THE HUMAN CYTOCHROME P450 2C
: TITLE OF INVENTION: SUBFAMILY
: NUMBER OF SEQUENCES: 44
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Townsend and Townsend Kourie and Crew
: STREET: 379 Lytton Avenue
: CITY: Palo Alto
: STATE: California
: COUNTRY: US
: ZIP: 94301
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC Compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/201,118
: FILING DATE: 22-FEB-1994
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/864,962
: FILING DATE: 09-APR-1992
: ATTORNEY/AGENT INFORMATION:
: NAME: Liebeschuetz, Joe
: REGISTRATION NUMBER: 37,505
: REFERENCE/DOCKET NUMBER: 15280-192-1
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (415) 326-2400
: TELEFAX: (415) 326-2422
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:

```

	FILING DATE:	06-MAY-1994
	CLASSIFICATION:	435
	PRIOR APPLICATION DATA:	
	APPLICATION NUMBER:	US 08/201,118
	FILING DATE:	22-FEB-1994
	CLASSIFICATION:	435
	PRIOR APPLICATION DATA:	
	APPLICATION NUMBER:	US 07/864,962
	FILING DATE:	09-APR-1992
	ATTORNEY/AGENT INFORMATION:	
	NAME:	Liebeschuetz, Joe
	REGISTRATION NUMBER:	37,505
	REFERENCE/DOCKET NUMBER:	15280-192110US
	TELEPHONE:	(650) 326-2400
	TELEFAX:	(650) 326-2422
	INFORMATION FOR SEQ ID NO:	1:
	SEQUENCE CHARACTERISTICS:	
	LENGTH:	490 amino acids
	TYPE:	amino acid
	TOPOLOGY:	unknown
	MOLECULE TYPE:	protein
	US-08-238-821B-1	
	Query Match	41.5%; Score 1084.5; DB 2; Length 490;
	Best Local Similarity	45.0%; Pred. No. 1.9e-106;
	Matches 217; Conservative	96; Mismatches 166; Indels 3; Gaps 3;
QY	9	LLALALLLLLTAALSGTRARGHLPGPPTPLPLGNLQLRLGALYSGLMRLSKKYGVPV 68
DB	7	LVLCLSLLLSIWROSS-GRGKLPPGPPTPLPVIGNILQIDIKDVSKSLTNLSKIYGVVF 65
QY	69	TIVLGPRWPVVVLVGDAVREALGGQAEEFSRGRTVMLEGTFDGHGVFFSNGERWRQR 128
DB	66	TLVFGLER-MVULHGVEVKELIDGEEFSGRHGFPLAERANRGFIVFSNGEKWKIR 124
QY	129	KFTMLALDLGMKGREGEELIQAEARCLVETFOGTGEGRPDPDSLLLAQATSNVVCSSLFG 188
DB	125	RFLMTLRNFMGKRSTEDRVQEARECLVEELKTKASPCDPTFILGCAPCNVICSIIFQ 184
QY	189	LRPSYEDKEQAVRAAGTLLGVSSOGGTYEMFSWFRLPLPGPHKOLLHHVSTLAAFT 248
DB	185	KRFYDKQQDFUNLMMEKNENIRIVSTPWIOICNNFTIIDYPFGTHNKLLKNAFWESDI 244
QY	249	VROVQOQHNLDSAGPARLDVADAFLLKMAOEQNPGTEFTNMKMLMTVIYLLFAGTMTVS 308
DB	245	LEKYKEHQSMDINN-RDFIDCFLIKWEKQKQOSEFTIENLVITAADLLGAGTTETS 303
QY	309	TTVGYTILLMKYPHVOKWRELNRGLGAGQAPSLGDRTRLPYTDAVLHEAQRLIALVP 368
DB	304	TLRYALLLLKHPEVTAKVOEIERVIGRRNSPCODRGHMPEYTDVAVHEVORYIDLIP 363
QY	369	MGIPRTLMTTRFRGYTLPGQTEVFPLLGSILSDPNIFKHPEEFNPDRFLDADGRFKHE 428
DB	364	TSLPHAVTCDEVKFNYLIPIKGTITLTSLVLDHNKEFPNPFDEPRHFDECGNFKXN 423
QY	429	AFLPFSGLKRCVLCGEGLAKELFFFTTILOAFLSPCPDPTLSLKPTVSGLFNIPPAP 488
DB	424	YFMPFSAKRICVGEGLARMELFLTILQNFNLKSLIDPKOLDTTPVVVNGFASVPPFY 483
QY	489	QL 490
DB	484	QL 485
	RESULT 13	
	PCT-US95-05744-1	
	Sequence 1, Application PC/TUS9505744	
	GENERAL INFORMATION:	
	APPLICANT:	GOLDSTEIN, Joyce A.
	APPLICANT:	ROMKES-SPARKS, Marjorie
	APPLICANT:	DE MORAIS, Sonia M.F.
	TITLE OF INVENTION:	CLONING, EXPRESSION AND DIAGNOSIS OF HUMAN
	TITLE OF INVENTION:	CYCLOCHROME P450 2C19: THE PRINCIPAL DETERMINANT OF S- </td
	TITLE OF INVENTION:	MEPHENTHOIN METABOLISM
	NUMBER OF SEQUENCES:	61
	CORRESPONDENCE ADDRESS:	
	ADDRESSEE:	Townsend and Crew LLP
	STREET:	Two Embarcadero Center, 8th Floor
	CITY:	San Francisco
	STATE:	California
	COUNTRY:	US
	ZIP:	94111
	COMPUTER READABLE FORM:	
	MEDIUM TYPE:	Floppy disk
	COMPUTER:	IBM PC compatible
	OPERATING SYSTEM:	PC-DOS/MS-DOS
	SOFTWARE:	PatentIn Release #1.0, Version #1.25
	CURRENT APPLICATION DATA:	
	FILE NAME:	US95-05744-1
	FILE DATE:	05/08/238 821B

;; TITLE OF INVENTION: CYTOCHROME P450 2C19: THE PRINCIPAL DETERMINANT  
;; NUMBER OF SEQUENCES: 61  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Townsend and Townsend Kourie and Crew  
;; STREET: 379 Lytton Avenue  
;; CITY: Palo Alto  
;; STATE: California  
;; COUNTRY: US  
;; ZIP: 94301  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: PCT/US95/05744  
;; FILING DATE:  
;; CLASSIFICATION:  
;; PRIOR APPLICATION NUMBER: US 08/238,821  
;; FILING DATE: 06-MAY-1994  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/201,118  
;; FILING DATE: 22-FEB-1994  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/864,962  
;; FILING DATE: 09-APR-1992  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Dow, Karen B.  
;; REGISTRATION NUMBER: 29,684  
;; REFERENCE/DOCKET NUMBER: 15280-192-1-1  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (415) 326-2400  
;; TELEFAX: (415) 326-2422  
;; INFORMATION FOR SEQ ID NO: 1:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 490 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: unknown  
;; MOLECULE TYPE: protein  
;; PCT-US95-05744-1  
  
Query Match 41.5%; Score 1084.5; DB 5; Length 490;  
Best Local Similarity 45.0%; Pred. No. 1.9e-106;  
Matches 217; Conservative 96; Mismatches 166; Indels 3; Gaps 3;  
  
QY 9 LLLALALLLLTALSSTRARGHLPPOGTPPLPLGNLLQLRPGALYSGLMRLSKYGPVF 68  
DB 7 LVLCCLLLLSLWRQSS-GRGKLPPOGTPPLPVIGNILQIDKDVSKSLNLSKIYGPVF 65  
QY 69 TIYLGWPWPVVLVQEAVERALGQAEFSRGRTVMAMLEGTDFGHGVFFSNGERWQLR 128  
DB 66 TLYFGLER-MVVLHGVEYVKEALIDGEEFSRGHFPPLAERANRGFIVFSNGKRWKEIR 124  
QY 129 KFTMLALDLGMGRGEGEELIQAEARCLVETFOGTEGRFPDPSLLLAQAATSNVVCSLIFG 188  
DB 125 RFSMLTLRNFMGKRSIEDRVQEAARCLVETFOGTEGRFPDPSLLLAQAATSNVVCSLIFQ 184  
QY 189 LRSYEDKEFQAVVRAAGTLLGVSSGGQGYEMFSWFLRPLGPHKQLLHVSTLAAFT 248  
DB 185 KRPDYKDOQFLNMEKLNENIRIVSTPWIQICNNFTIIDYFPGTHKNLLKNAFMSDI 244  
QY 249 VROVQOQGNLDASGPARDLVDAFLKMAQEEQNGPTEFTNKMMLTVIYLLFAGTWTVS 308  
DB 245 LEVQKHQESMDINNP-RDFIDFLNMEKEKQSQSEFTIENLVITAADLLGAGTETS 303  
QY 309 TTVGYTLLLMKYPHYQKQVREELNRELGAQAPSLGDRTRLPYTDVAVLHEAQRLLALYP 368  
DB 304 TTLRYALLLLKHPEYAKVQOEIERNVIGNRSPCQDRGHMPTTDAVVEVQRYIDLIP 363  
QY 369 MGLPRTLMTRTRPGTYTLFQGTVEVFFLLGSILHDPNIFKHPEEFNPDRLDADGRFKHE 428

DB 364 TSLPHAVTCDVKFRNYLIPKGTITLTSVLDHNKEPPNEMFDRHFLDEGGNPKSN 423  
QY 429 AFLPESLGRKVCGLGELAKAELFFETTLTQAFSLESPPDPTLSLKPTVSGLFNIPPAF 488  
DB 424 YFMPSAGKRICVGBGLARMELFLFTLILQNFNLKSLIDPKDLDTTPVWNGFASVPPFY 483  
QY 489 QL 490  
DB 484 QL 485  
  
RESULT 14  
US-08-201-118-5  
; Sequence 5, Application US/08201118  
; Patent No. 5786191  
; GENERAL INFORMATION:  
; APPLICANT: GOLDSTEIN, Joyce A.  
; TITLE OF INVENTION: CLONING AND EXPRESSION OF COMPLEMENTARY  
; TITLE OF INVENTION: DNAs FOR MULTIPLE MEMBERS OF THE HUMAN CYTOCHROME P450 2C  
; NUMBER OF SEQUENCES: 44  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend Kourie and Crew  
; STREET: 379 Lytton Avenue  
; CITY: Palo Alto  
; STATE: California  
; COUNTRY: US  
; ZIP: 94301  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/201,118  
; FILING DATE: 22-FEB-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/864,962  
; FILING DATE: 09-APR-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Liebeschuetz, Joe  
; REGISTRATION NUMBER: 37,505  
; REFERENCE/DOCKET NUMBER: 15280-192-1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 326-2400  
; TELEFAX: (415) 326-2422  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 490 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-201-118-5  
  
Query Match 41.1%; Score 1073.5; DB 1; Length 490;  
Best Local Similarity 43.9%; Pred. No. 2.7e-105;  
Matches 212; Conservative 97; Mismatches 171; Indels 3; Gaps 3;  
  
QY 8 ALLALALLLLTALSSTRARGHLPPOGTPPLPLGNLLQLRPGALYSGLMRLSKYGPV 67  
DB 6 ALVLCCLLLLSLWRQSS-GRGKLPPOGTPPLPVIGNILQIDKDVSKSLNLSKIYGPV 64  
QY 68 TIYLGWPWPVVLVQEAVERALGQAEFSRGRTVMAMLEGTDFGHGVFFSNGERWQL 127  
DB 65 FTIVYFG-LKPTVVLHGVEYVKEALIDHGEFSGRGSFPVAEKVKNKGLILFSGKRWKEI 123  
QY 128 RKFTMLALDLGMGRGEGEELIQAEARCLVETFOGTEGRFPDPSLLLAQAATSNVVCSLIF 187  
DB 124 RFLCLMTLRFNMGKRSIEDRVQEAARCLVETFOGTEGRFPDPSLLLAQAATSNVVCSLIF 183



QY 188 GLRFSYEDKEFOAVVRAAGGTLGVSSOGGQTYEMFSWFLRPLPGPHKQLLHHVSTLAAF 247  
Db 184 HDRFDYKQDFLNLMEKFNENLRILSSPWIQVCNNFPALIDYLPQSHNKIAENFAIYIKSY 243  
QY 248 TVRQVOOQGNLDASGPARDLVDAFLKMAQBEQNPGEFTNKNMLMTVIVILLFAGTMTV 307  
Db 244 VLRIKEHOESLDMS-ARDFIDCFLLKMEQKHQOQSEFTVESLIAVTDMEFGAGTETT 302  
QY 308 STTVGYTLLLMKYPHVQKWRVREELNRELGAQAPSLGDRTRLPYTDVAVLHQAORLLALV 367  
Db 303 STTLRYGLLLLLKYPEVTAKVQEEIECVVGRNRSPCMQDRSHMPYTDVAVHEIQRYIDL 362  
QY 368 PMGIPRTLMRTTRFRGYTLPGQTEVFPPLGSLTHDPNFKHPEEFNPDRLDADGRFRKH 427  
Db 363 PTNLPHAVTCDVKFKNYLIPKGTITITSLSVLHNDKEFPNEMFDPGHFLDKSGNFKKS 422  
QY 428 EAPLPSLOKRVCLGKGLAKAELFLFPTTILQAFSLGSPCPDPTLSLKPVTSGLFNIPPA 487  
Db 423 DYFMFPSAGKRCMCGELARMELFLFTTILQNFNLKSQVDPKIDITPIANAFGRVPL 482  
QY 488 FOL 490  
Db 483 YQL 485

RESULT 15  
US-08-238-821B-5  
; Sequence 5, Application US/08238821B  
; Patent No. 5912120  
; GENERAL INFORMATION:  
; APPLICANT: GOLDSTEIN, Joyce A.  
; APPLICANT: ROMKES-SPARKS, Marjorie  
; APPLICANT: DE MORAIS, Sonia M.F.  
; TITLE OF INVENTION: CLONING, EXPRESSION AND DIAGNOSIS OF HUMAN  
; TITLE OF INVENTION: CYTOCHROME P450 2C19: THE PRINCIPAL DETERMINANT OF S-  
; TITLE OF INVENTION: MEPHENYTOIN METABOLISM  
; NUMBER OF SEQUENCES: 61  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, 8th Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: US  
; ZIP: 94111  
; COMPUTER READABLE FORM: disk  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/238,821B  
; FILING DATE: 06-MAY-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/201,118  
; FILING DATE: 22-FEB-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/864,962  
; FILING DATE: 09-APR-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Liebeschuetz, Joe  
; REGISTRATION NUMBER: 37,505  
; REFERENCE/DOCKET NUMBER: 15280-192110US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (650) 326-2400  
; TELEFAX: (650) 326-2422  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 490 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear

i MOLECULE TYPE: protein  
US-08-238-821B-5  
Query Match 41.1%; Score 1073.5; DB 2; Length 490;  
Best Local Similarity 43.9%; Pred. No. 2.7e-105;  
Matches 212; Conservative 97; Mismatches 171; Indels 3; Gaps 3;  
QY 8 ALLLALALLLLTALSGTRAGHLPGPPTPLPLIAGNLLQLRFGALYSGIAMLRLSKYGV 67  
Db 6 ALVLCCLCLFLLSLWRQSS-GRGRLPSPGPTPLPIIGNILQLDVKDKSKSLTNFSKYGV 64  
QY 68 FTIYLGPNRPVVVLVGQAEAREALGGQABEFSRGRTGVTAMLEGTDPGHGVFFNGSRWQL 127  
Db 65 FTVYFG-LKPIVHLHGVEAVKEALIDHGEFEFSGRGSFFVAEKVKNKGLGILFSGKRWKEI 123  
QY 128 RKFTMLARDLGMGKEGEELIQAERCLVETFOCTEGRPDPSPSLLLAQATSNVCSLLF 187  
Db 124 RRFCLMTLURNFGMGKRSIEDRQVEARCLVEBLRKTNASPCDPTFILGAPCNVICSVIF 183  
QY 188 GLRFSYEDKEFOAVVRAAGGTLGVSSOGGQTYEMFSWFLRPLPGPHKQLLHHVSTLAAF 247  
Db 184 HDRFDYKQDFLNLMEKFNENLRILSSPWIQVCNNFPALIDYLPQSHNKIAENFAIYIKSY 243  
QY 248 TVRQVOOQGNLDASGPARDLVDAFLKMAQBEQNPGEFTNKNMLMTVIVILLFAGTMTV 307  
Db 244 VLRIKEHOESLDMS-ARDFIDCFLLKMEQKHQOQSEFTVESLIAVTDMEFGAGTETT 302  
QY 308 STTVGYTLLLMKYPHVQKWRVREELNRELGAQAPSLGDRTRLPYTDVAVLHQAORLLALV 367  
Db 303 STTLRYGLLLLLKYPEVTAKVQEEIECVVGRNRSPCMQDRSHMPYTDVAVHEIQRYIDL 362  
QY 368 PMGIPRTLMRTTRFRGYTLPGQTEVFPPLGSLTHDPNFKHPEEFNPDRLDADGRFRKH 427  
Db 363 PTNLPHAVTCDVKFKNYLIPKGTITITSLSVLHNDKEFPNEMFDPGHFLDKSGNFKKS 422  
QY 428 EAPLPSLOKRVCLGKGLAKAELFLFPTTILQAFSLGSPCPDPTLSLKPVTSGLFNIPPA 487  
Db 423 DYFMFPSAGKRCMCGELARMELFLFTTILQNFNLKSQVDPKIDITPIANAFGRVPL 482  
QY 488 FOL 490  
Db 483 YQL 485

Search completed: September 15, 2004, 09:29:41  
Job time : 35 secs

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2004, 09:14:35 ; Search time 128 Seconds

(without alignments)  
1112.531 Million cell updates/sec

Title: US-10-669-693-2

Perfect score: 2615

Sequence: 1 MEATGTWALLALALLLLT.....PPAFQLQVRPTDLHSTTQTR 504

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq\_29Jan04:\*

1: Geneseqp1980s:\*

2: Geneseqp1990s:\*

3: Geneseqp2000s:\*

4: Geneseqp2001s:\*

5: Geneseqp2002s:\*

6: Geneseqp2003as:\*

7: Geneseqp2003bs:\*

8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2615	100.0	504	4 AAU12203	Aau12203 Human PRO
2	2615	100.0	504	5 AAE28282	Aae28282 Human PRO
3	2615	100.0	504	5 AAE15439	Aae15439 Human PRO
4	2615	100.0	504	6 ABO17647	Abol17647 Novel hum
5	2615	100.0	504	6 ABU10222	Abu10222 Human cyt
6	2615	100.0	504	6 ABU80901	Abu80901 Human PRO
7	2615	100.0	504	6 ABU66601	Abu66601 Human PRO
8	2615	100.0	504	6 ABUS9682	Abu59682 Novel sec
9	2615	100.0	504	6 ABUS9682	Abu59682 Human sec
10	2615	100.0	504	6 ABO24872	Abo24872 Human sec
11	2615	100.0	504	6 ABUF6877	Abu6877 Human sec
12	2615	100.0	504	6 ADA45583	Ada45583 Human PRO
13	2615	100.0	504	6 ADA76014	Ada76014 Human PRO
14	2615	100.0	504	6 ADA18664	Ada18664 Human PRO
15	2615	100.0	504	6 ADA61287	Ada61287 Homo sapi
16	2615	100.0	504	6 ADB19072	Adb19072 Novel hum
17	2615	100.0	504	6 ADB27613	Adb27613 Human PRO
18	2615	100.0	504	6 ADB86092	Adb86092 Novel hum
19	2615	100.0	504	6 ADB15656	Adb15656 Human PRO
20	2615	100.0	504	6 ADA47442	Ada47442 Human PRO
21	2615	100.0	504	6 ADA67237	Ada67237 Human PRO
22	2615	100.0	504	6 ADB30244	Adb30244 Human PRO
23	2615	100.0	504	6 ADA85540	Ada85540 Novel hum
24	2615	100.0	504	6 ADA96752	Ada96752 Human PRO
25	2615	100.0	504	6 ADA79056	Ada79056 Human PRO
				6 ADA87195	Ada87195 Novel hum

#### ALIGNMENTS

##### RESULT 1

AAU12203

ID AAU12203 standard; protein; 504 AA.

AC AAU12203;

XX 24-OCT-2001 (first entry)

DT Human PRO1906 polypeptide sequence.

DE Human secretory and transmembrane; PRO; mammalian; cancer; lung; breast;

KW prostate; cervical; tumour necrosis factor-alpha; TNF-alpha; cartilage;

KW ear; proliferation; glucose; free fatty acid; skeletal muscle; adipocyte;

KW A-peptide; factor VIIA; gene therapy.

OS Homo sapiens.

XX WO200140466-A2.

PN 07-JUN-2001.

PD 01-DEC-2000; 2000WO-US032678.

PF 01-DEC-1999; 99WO-US028301.

XX 01-DEC-1999; 99WO-US028634.

PR 02-DEC-1999; 99WO-US028551.

PR 02-DEC-1999; 99WO-US028564.

PR 02-DEC-1999; 99WO-US028565.

PR 05-DEC-1999; 99US-0170262P.

PR 16-DEC-1999; 99WO-US030095.

PR 20-DEC-1999; 99WO-US030911.

PR 30-DEC-1999; 99WO-US031243.

PR 30-DEC-1999; 99WO-US031274.

PR 05-JAN-2000; 2000WO-US000219.

PR 06-JAN-2000; 2000WO-US000277.

PR 11-FEB-2000; 2000WO-US000376.

PR 18-FEB-2000; 2000WO-US003565.

PR 18-FEB-2000; 2000WO-US004341.

PR 22-FEB-2000; 2000WO-US004342.

PR 24-FEB-2000; 2000WO-US004914.

PR 24-FEB-2000; 2000WO-US004914.

PR 01-MAR-2000; 2000WO-US005004.

PR 02-MAR-2000; 2000WO-US005841.

PR 03-MAR-2000; 2000US-0187202P.

PR 10-MAR-2000; 2000WO-US006319.

PR 15-MAR-2000; 2000WO-US006884.

ADb16397 Human PRO  
Ada91489 Novel hum  
ADb14552 Human PRO  
ADb18513 Novel hum  
Ada93728 Human PRO  
ADb19624 Novel hum  
ADb12936 Human PRO  
Abo43180 Novel hum  
Ada74130 Human PRO  
ADb24423 Human PRO  
Ada81947 Human PRO  
Ada74910 Human PRO  
Ada84988 Novel hum  
Ada84436 Novel hum  
ADb29692 Human PRO  
Ada80220 Human PRO  
Ada75462 Human PRO  
Ada46687 Human PRO  
ADb24983 Human PRO  
Ada93159 Human PRO

PR	20-MAR-2000;	2000WO-US007377.	
PR	21-MAR-2000;	2000WO-US007532.	
PR	30-MAR-2000;	2000WO-US008439.	
PR	17-MAY-2000;	2000WO-US013705.	
PR	22-MAY-2000;	2000WO-US014042.	
PR	30-MAY-2000;	2000WO-US014941.	
PR	02-JUN-2000;	2000WO-US015264.	
PR	05-JUN-2000;	2000US-0209832P.	
PR	28-JUN-2000;	2000WO-US020710.	
PR	11-AUG-2000;	2000WO-US022031.	
PR	23-AUG-2000;	2000WO-US023522.	
PR	24-AUG-2000;	2000WO-US023328.	
PR	08-NOV-2000;	2000WO-US030952.	
PR	10-NOV-2000;	2000WO-US030873.	
XX		(GETH ) GENENTECH INC.	
PA			
XX	Baker KP, Beresini M, Deforge L, Desnoyers L, Filvaroff E, Gao W;		
PI	Gerritsen ME, Goddard A, Godowski FJ, Gurney AL, Sherwood S;		
PI	Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;		
XX			
DR	WPI; 2001-408281/43.		
DR	N-PSDB; AAE21275.		
XX			
XX	Isolated , secretory and transmembrane PRO polypeptide used to detect		
XX	other PRO polypeptides, link bioactive molecules to cells expressing PRO		
PT	polypeptides, and detect the presence of mammalian tumors e.g. lung,		
PT	breast, prostate, cervical.		
PT			
XX			
XX	Claim 12; Fig 64; 813pp; English.		
XX			
CC	AAU12172-AAU12446 represent novel human secretory and transmembrane PRO		
CC	polypeptides. The PRO polypeptides are useful to detect other PRO		
CC	polypeptides, to link bioactive molecules to cells expressing PRO		
CC	polypeptides, to modulate biological activities of cells expressing PRO		
CC	polypeptides, and to detect the presence of mammalian lung, colon,		
CC	breast, prostate, rectal, cervical or liver tumours by comparing PRO		
CC	polypeptide expression in a cell sample to that in a control sample. Some		
CC	of the 275 sequences are also useful to stimulate the release of tumour		
CC	necrosis factor-alpha (TNF-alpha) from human blood, the proliferation or		
CC	differentiation of chondrocytes, the proliferation or gene expression in		
CC	pericyte cells, the release of proteoglycans from cartilage, the		
CC	proliferation of inner ear utricular supporting cells or of T-		
CC	lymphocytes, the release of a cytokine from peripheral blood monocytes		
CC	(PBMCs), or the proliferation of endothelial cells. Some of the PRO		
CC	polypeptides may modulate glucose or free fatty acid uptake by skeletal		
CC	muscle cells or by adipocytes; or inhibit binding of A-peptide to factor		
CC	VIIA. The PRO polypeptides can be used in assays to identify molecules		
CC	involved in binding interactions. The polynucleotides encoding PRO		
CC	polypeptides can be used to generate probes, antisense RNA/DNA,		
CC	transgenic or knock out animals and can be used in gene therapy		
XX			
SQ	Sequence 504 AA;		
	Query Match 100.0%; Score 2615; DB 4; Length 504;		
	Best Local Similarity 100.0%; Pred. No. 1.8e-258;		
	Matches 504; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
Qy	1 MEATGTWALLLALALLLLTLLSGTRARGHLPPTPLLLGNLLQRLPGALYGLMRL 60		
Db	1 MEATGTWALLLALALLLLTLLSGTRARGHLPPTPLLLGNLLQRLPGALYGLMRL 60		
Qy	61 SKKYGVPFTIYLGWPWPVVVLVGGQAVREALGGQAEFFSGRTGVTAMLEGTPDHGVFFSN 120		
Db	61 SKKYGVPFTIYLGWPWPVVVLVGGQAVREALGGQAEFFSGRTGVTAMLEGTPDHGVFFSN 120		
Qy	121 GERWRQLKFTMLALRDLCMKRGEELIQAEARCLVETFGTEGRPPDPSLLLAQATSN 180		
Db	121 GERWRQLKFTMLALRDLCMKRGEELIQAEARCLVETFGTEGRPPDPSLLLAQATSN 180		
Qy	181 VVCSLLFLGRFSYEDKEFOAVVRAAGTLLGVSSGGQTYEMFSLRPLFCPHKQLLHH 240		
Db	181 VVCSLLFLGRFSYEDKEFOAVVRAAGTLLGVSSGGQTYEMFSLRPLFCPHKQLLHH 240		
Qy	241 VSTLAFTVRQVQHQGNLDASGPARDLVDAFLLLKMAQEEQNPCTFTNKMLMTVIYLL 300		
Db	241 VSTLAFTVRQVQHQGNLDASGPARDLVDAFLLLKMAQEEQNPCTFTNKMLMTVIYLL 300		
Qy	301 FAGTMTVSTTVGYTLLLLMKYPHVQKWVREELNRELGAQAPSLGDRTRLEYTDAVLHEA 360		
Db	301 FAGTMTVSTTVGYTLLLLMKYPHVQKWVREELNRELGAQAPSLGDRTRLEYTDAVLHEA 360		
Qy	361 QRLIALVPMGIPTLMRTTRRGYTLPGTVEFPLGSLILHDPNIFKHPEEFNPDRLDA 420		
Db	361 QRLIALVPMGIPTLMRTTRRGYTLPGTVEFPLGSLILHDPNIFKHPEEFNPDRLDA 420		
Qy	421 DGRPKHEAFPLPFSGLKRVCLGEGAKAEFLFFTTILQAFLSLESCPPDRLSLKPTVSG 480		
Db	421 DGRPKHEAFPLPFSGLKRVCLGEGAKAEFLFFTTILQAFLSLESCPPDRLSLKPTVSG 480		
Qy	481 LFNIPPAFQLQVRPTDLHSTTQTR 504		
Db	481 LFNIPPAFQLQVRPTDLHSTTQTR 504		
XX			
XX	AAE28282		
ID	AAE28282 standard; protein; 504 AA.		
XX	AAE28282;		
XX	27-DEC-2002 (first entry)		
XX	Human drug-metabolising enzyme.		
XX	Human; drug-metabolising enzyme; cytochrome p450; therapy; cytostatic;		
XX	Parkinson's disease; myotonic dystrophy; developmental defect; enzyme;		
XX	cancer; chromosome 19.		
XX	Homo sapiens.		
XX	Key Location/Qualifiers		
FT	Region 5..25 Helix 1		
FT	Region 32..53 Helix 2		
FT	Region 39..60 Helix 3		
FT	Modified-site 51..56 Leucine zipper pattern		
FT	Modified-site 61..63 N-myristoylation site		
FT	Region 64..84 Protein kinase C phosphorylation site		
FT	Modified-site 99..101 Helix 2		
FT	Modified-site 109..114 Protein kinase C phosphorylation site		
FT	Modified-site 115..120 N-myristoylation site		
FT	Modified-site 119..122 N-myristoylation site		
FT	Modified-site 128..131 Casein kinase II phosphorylation site		
FT	phosphorylation site		
FT	140..143 CAMP- and cGMP-dependent protein kinase		
FT	170..190 Amidation site		
FT	188..193 Helix 3		
FT	Modified-site 192..195 N-myristoylation site		
FT	Modified-site 207..212 Casein kinase II phosphorylation site		
FT	Modified-site 248..250 N-myristoylation site		

FT Modified-site /note= "Protein kinase C phosphorylation site"  
 FT 257..262  
 FT /note= "N-myristoylation site"  
 FT 284..289  
 FT /note= "N-myristoylation site"  
 FT 288..290  
 FT /note= "Protein kinase C phosphorylation site"  
 FT 292..312  
 FT /note= "Helix 4"  
 FT 339..344  
 FT /note= "N-myristoylation site"  
 FT 343..346  
 FT /note= "Casein kinase II phosphorylation site"  
 FT 370..375  
 FT /note= "N-myristoylation site"  
 FT 378..380  
 FT /note= "Protein kinase C phosphorylation site"  
 FT 435..438  
 FT /note= "Amidation site"  
 FT 444..449  
 FT /note= "N-myristoylation site"  
 FT 448..468  
 FT /note= "Helix 5"  
 FT 473..475  
 FT /note= "Protein kinase C phosphorylation site"  
 FT XX US2002076774-A1.  
 FN 20-JUN-2002.  
 PD 27-DEC-2000; 2000US-00748127.  
 PF 21-JUN-2000; 2000US-0212840P.  
 XX (YANC/) YAN C.  
 PA (SHAO/) SHAO W.  
 PA (MERK/) MERKULOV G V.  
 PA (DFRA/) DI FRANCESCO V.  
 PA (BEAS/) BEASLEY E M.  
 XX Yan C, Shao W, Merkulov GV, Di Francesco V, Beasley EM;  
 PI WPI; 2002-673347/72.  
 XX N-PSDB; AAD45438, AAD45439.  
 DR A new human drug-metabolizing enzyme is related to the cytochrome p450  
 XX superfamily and is useful to provide modulators to treat drug-  
 PT metabolizing enzyme-related disorders including cancer and Parkinson's  
 PT disease.  
 XX Claim 1; Page 22-23; 49pp; English.  
 XX The invention relates to human drug-metabolising enzyme peptides related  
 CC to the cytochrome p450 superfamily. Binding agents of the peptide are  
 CC used to treat a disease or condition mediated by a human drug-  
 CC metabolising enzyme. Such diseases include Parkinson's disease, cancer,  
 CC myotonic dystrophy and developmental defects. The present sequence is  
 CC human drug-metabolising enzyme. The gene encoding this enzyme is located  
 CC at chromosome 19  
 XX Sequence 504 AA;  
 SQ  
 Query Match 100.0%; Score 2615; DB 5; Length 504;  
 Best Local Similarity 100.0%; Pred. No. 1.8e-258;  
 Matches 504; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MEATGTWALLALALLLLTALSGTRARGHLPPGPTPLPLGNLLQRLPGALYGLMRL 60  
 DB 1 MEATGTWALLALALLLLTALSGTRARGHLPPGPTPLPLGNLLQRLPGALYGLMRL 60  
 QY 61 SKKYGPFVITYLGPWPVVVLVQEAIVREAGLQGAEEFSGRGTVAMLEGTFDGHGVFFSN 120  
 DB 61 SKKYGPFVITYLGPWPVVVLVQEAIVREAGLQGAEEFSGRGTVAMLEGTFDGHGVFFSN 120

QY 121 GERWQLRKFTMLALRDLMGKREGEELIQAEARCLIVETFOGTGEPFDPSPSLLAQATSN 180  
 DB 121 GERWQLRKFTMLALRDLMGKREGEELIQAEARCLIVETFOGTGEPFDPSPSLLAQATSN 180  
 QY 181 VVCSLLFLGRFSYEDKEFAQAVVRAAGTLLGVSSQGGQTYEMPSWFLRPLPGPHKOLLHH 240  
 DB 181 VVCSLLFLGRFSYEDKEFAQAVVRAAGTLLGVSSQGGQTYEMPSWFLRPLPGPHKOLLHH 240  
 QY 241 VSTLAFTVROVQOQGNLDASGPARDLVDAFLKMAQEQNPQTEFTKNMLMTVIYLL 300  
 DB 241 VSTLAFTVROVQOQGNLDASGPARDLVDAFLKMAQEQNPQTEFTKNMLMTVIYLL 300  
 QY 301 FAGTMTVSTTVGYTLLLLMKYPHVQKVRREELNRELGAQAPSLGDRTRLPYTDALVHEA 360  
 DB 301 FAGTMTVSTTVGYTLLLLMKYPHVQKVRREELNRELGAQAPSLGDRTRLPYTDALVHEA 360  
 QY 361 QRLALVPMGIPRTLMTTRFRGYTLPGQTEVFPFLLGSLIHDHPNFKHPEEPNDRFLDA 420  
 DB 361 QRLALVPMGIPRTLMTTRFRGYTLPGQTEVFPFLLGSLIHDHPNFKHPEEPNDRFLDA 420  
 QY 421 DGRFRKHEAFLPFSLGKRVCLGEGLAKELELFFTTILOAFSLSPCPDPTLSLKPTVSG 480  
 DB 421 DGRFRKHEAFLPFSLGKRVCLGEGLAKELELFFTTILOAFSLSPCPDPTLSLKPTVSG 480  
 QY 481 LFNIPPAFQLQVRPTDLHSTTQTR 504  
 DB 481 LFNIPPAFQLQVRPTDLHSTTQTR 504  
 RESULT 3  
 ID AAE15439 standard; protein; 504 AA.  
 XX AAE15439;  
 XX 12-MAR-2002 (first entry)  
 XX Human drug metabolising enzyme (DME)-6.  
 DE Human; drug metabolising enzyme; gene therapy; autoimmune disorder;  
 KW inflammatory disorder; acquired immune deficiency syndrome; infection;  
 KW AIDS; adult respiratory distress syndrome; cell proliferative disorder;  
 KW allergy; anaemia; conjunctivitis; actinic keratosis; arteriosclerosis;  
 KW cancer; endocrine disorder; hypothalamus disorder; pituitary disorder;  
 KW gastrointestinal disorder; metabolic disorder; developmental disorder;  
 KW liver disorder; iritis; cystic fibrosis; Addison's disease; retinitis;  
 KW goitre; diabetes; dysphagia; gastric carcinoma; anorexia; eye disorder;  
 KW DME-6.  
 XX Homo sapiens.  
 OS  
 FH Key Location/Qualifiers  
 FT Peptide 1..28  
 FT /label= Signal\_peptide  
 FT Domain 7..26  
 FT /note= "Transmembrane domain"  
 FT Protein 29..504  
 FT /note= "Human mature DME-6 protein"  
 FT Domain 297..315  
 FT /note= "Transmembrane domain"  
 XX WO200179468-A2.  
 PN 25-OCT-2001.  
 XX 12-APR-2001; 2001WO-US011869.  
 PF 13-APR-2000; 2000US-0197590P.  
 XX 19-APR-2000; 2000US-0198403P.  
 PR 28-APR-2000; 2000US-0200185P.  
 PR 05-MAY-2000; 2000US-0202234P.  
 PR 11-MAY-2000; 2000US-0203509P.



PR	30-NOV-1999;	99WO-US028313.
PR	30-NOV-1999;	99WO-US028409.
PR	01-DEC-1999;	99WO-US028301.
PR	01-DEC-1999;	99WO-US028634.
PR	02-DEC-1999;	99WO-US028551.
PR	02-DEC-1999;	99WO-US028564.
PR	02-DEC-1999;	99WO-US028565.
PR	16-DEC-1999;	99WO-US030095.
PR	20-DEC-1999;	99WO-US030911.
PR	20-DEC-1999;	99WO-US030999.
PR	22-DEC-1999;	99WO-US030720.
PR	30-DEC-1999;	99WO-US031243.
PR	30-DEC-1999;	99WO-US031274.
PR	05-JAN-2000;	2000WO-US000217.
PR	06-JAN-2000;	2000WO-US000279.
PR	11-FEB-2000;	2000WO-US000376.
PR	11-FEB-2000;	2000WO-US003365.
PR	18-FEB-2000;	2000WO-US004341.
PR	18-FEB-2000;	2000WO-US004342.
PR	22-FEB-2000;	2000WO-US004414.
PR	24-FEB-2000;	2000WO-US004914.
PR	24-FEB-2000;	2000WO-US005004.
PR	01-MAR-2000;	2000WO-US005601.
PR	02-MAR-2000;	2000WO-US005746.
PR	10-MAR-2000;	2000WO-US005841.
PR	10-MAR-2000;	2000WO-US006319.
PR	15-MAR-2000;	2000WO-US006884.
PR	21-MAR-2000;	2000WO-US007377.
PR	21-MAR-2000;	2000WO-US007532.
PR	17-MAY-2000;	2000WO-US008439.
PR	22-MAY-2000;	2000WO-US013705.
PR	30-MAY-2000;	2000WO-US014042.
PR	02-JUN-2000;	2000WO-US014941.
PR	02-JUN-2000;	2000WO-US015264.
PR	11-AUG-2000;	2000WO-US020710.
PR	23-AUG-2000;	2000WO-US022031.
PR	24-AUG-2000;	2000WO-US023322.
PR	08-NOV-2000;	2000WO-US030952.
PR	10-NOV-2000;	2000WO-US030873.
PR	01-DEC-2000;	2000WO-US032678.
PR	20-DEC-2000;	2000US-US0074259.
PR	20-DEC-2000;	2000US-US0074259.
PR	28-FEB-2001;	2001US-US00796498.
PR	28-FEB-2001;	2001US-US00796498.
PR	01-MAR-2001;	2001WO-US006520.
PR	09-MAR-2001;	2001WO-US006566.
PR	14-MAR-2001;	2001US-US0802706.
PR	22-MAR-2001;	2001US-US080689.
PR	25-MAR-2001;	2001US-US0816744.
PR	05-APR-2001;	2001US-US082366.
PR	10-MAY-2001;	2001US-US084208.
PR	10-MAY-2001;	2001US-US084508.
PR	18-MAY-2001;	2001US-US0860216.
PR	25-MAY-2001;	2001US-US086028.
PR	25-MAY-2001;	2001US-US0866034.
PR	01-JUN-2001;	2001WO-US017092.
PR	01-JUN-2001;	2001US-US0872035.
PR	21-JUN-2001;	2001US-US0887879.
PR	21-JUN-2001;	2001WO-US020116.
PR	09-JUL-2001;	2001WO-US021066.
PR	09-JUL-2001;	2001WO-US021735.
PR	18-JUL-2001;	2001US-US090827.
PR	06-AUG-2001;	2001US-US0924419.
PR	09-AUG-2001;	2001US-US0927796.
PR	16-AUG-2001;	2001US-US0931836.
PR	19-DEC-2001;	2001US-US028072.

Baker KP, Beresini M, Deforge L, Desnoyers L, Filvaroff E, Gao W;  
Gerritsen ME, Goddard A, Godowski RJ, Gurney AL, Sherwood S;  
Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;  
WPI: 2003-341980/32.  
N-PSDB; ACD23884.

New secreted and transmembrane PRO nucleic acids, for treating  
inflammation, organ failure, atherosclerosis, cardiac injury,  
infertility, birth defects, premature aging, acquired immunodeficiency  
syndrome (AIDS), or cancer.

Claim 12; Fig 64; 660pp; English.

The invention describes an isolated nucleic acid (I) comprising, or which  
has 80 % sequence identity to, or the full-length coding sequence of, one  
of 275 nucleotide sequences, and which encodes a corresponding  
polypeptide selected from 275 amino acid sequences, where all sequences  
are given in the specification. The polypeptide encoded by (I) is used to  
detect PRO polypeptides, link a bioactive molecule to a cell expressing a  
PRO polypeptide, modulate a biological activity of a cell, stimulate the  
release of tumour necrosis factor (TNF)-alpha from human blood, modulate  
the uptake of glucose or free fatty acid by cells, stimulate or inhibit  
the proliferation or differentiation of cells or gene expression,  
stimulate the release of proteoglycans, stimulate the release of cytokine  
from peripheral blood mononuclear cells, inhibit the binding of A-peptide  
to factor VIIA, or detect the presence of tumour in a mammal. The nucleic  
acid and polypeptide encoded by it, are useful for treating inflammatory  
diseases, organ failure, atherosclerosis, cardiac injury, infertility,  
birth defects, premature aging, acquired immunodeficiency syndrome  
(AIDS), cancer, or diabetic complications. The nucleic acid is useful as  
hybridisation probes, in chromosome and gene mapping, and in generating  
antisense RNA or DNA. The polypeptides are useful as pharmaceuticals,  
diagnostics, biosensors or bioreactors. Both are useful in tissue typing.  
This is the amino acid sequence of a novel human secreted and  
transmembrane PRO polypeptide

Sequence 504 AA:  
SQ

```
Db      421  DGRFRKHEAFPSLGRKVCIGEGIAKAEFLFFTTILQAFSLSPCPDPTLSLKPTVSG 480
QY      481  LFNIPPAFQLOVRPTDLHSTTQTR 504
Db      481  LFNIPPAFQLOVRPTDLHSTTQTR 504

RESULT 5
ABU10222
ID  ABU10222 standard; protein; 504 AA.
XX
AC  ABU10222;
XX
DT  28-JUL-2003 (first entry)
XX
DE  Human cytochrome P450, 33303.
XX
KW  Human; cytochrome; P450; cellular proliferation; cancer; migraine;
KW  cellular differentiation; cell adhesion; cell motility; cell migration;
KW  inflammation; autoimmune disorder; rheumatoid arthritis; diabetes;
KW  multiple sclerosis; systemic lupus erythematosus; conjunctivitis; asthma;
KW  allergy; Grave's disease; graft-versus-host disease; vascular disorder;
KW  central nervous system disorder; Alzheimer's disease; anxiety disorder;
KW  amnesia; depression; Parkinson's disease; obesity; cachexia; anorexia;
KW  cell-signaling-associated disorder; metabolism-associated disorder;
KW  steroid-associated disorder; fatty acid-associated disorder.
XX
OS  Homo sapiens.
XX
FH  Key
FT  Peptide
FT  1..29 Location/Qualifiers
FT  30..504 /label= Signal_sequence
FT  /label= Mature_33303
FT  32..53 Region
FT  /label= Leucine_zipper
FT  33..493 Domain
FT  /label= Cytochrome_P450_domain
FT  51..56 Modified-site
FT  /note= "N-myristoylation site"
FT  61..63 Modified-site
FT  /note= "Protein kinase C phosphorylation site"
FT  99..102 Binding-site
FT  /label= Glycosaminoglycan_attachment_site
FT  99..101 Modified-site
FT  /note= "Protein kinase C phosphorylation site"
FT  109..114 Modified-site
FT  /note= "N-myristoylation site"
FT  115..120 Modified-site
FT  /note= "N-myristoylation site"
FT  119..122 Modified-site
FT  /note= "Casein kinase II phosphorylation site"
FT  128..131 Modified-site
FT  /note= "cAMP and cGMP-dependent protein kinase phosphorylation site"
FT  140..143 Modified-site
FT  /note= "Amidation site"
FT  188..193 Modified-site
FT  /note= "N-myristoylation site"
FT  192..195 Modified-site
FT  /note= "Casein kinase II phosphorylation site"
FT  207..212 Modified-site
FT  /note= "N-myristoylation site"
FT  248..250 Modified-site
FT  /note= "Protein kinase C phosphorylation site"
FT  257..261 Modified-site
FT  /note= "N-myristoylation site"
FT  284..289 Modified-site
FT  /note= "N-myristoylation site"
FT  288..290 Modified-site
FT  /note= "Protein kinase C phosphorylation site"
FT  339..344 Modified-site
```

```
FT Modified-site /note= "N-myristoylation site"
FT 343..346 /note= "Casein kinase II phosphorylation site"
FT Modified-site /note= "N-myristoylation site"
FT 370..375 /note= "N-myristoylation site"
FT Modified-site /note= "Protein kinase C phosphorylation site"
FT 378..380 /note= "Protein kinase C phosphorylation site"
FT Region 433..442 /note= "Cytochrome P450 cysteine haem-iron ligand signature"
FT Modified-site 435..438 /note= "Amidation site"
FT Modified-site 444..449 /note= "N-myristoylation site"
FT Modified-site 473..475 /note= "Protein kinase C phosphorylation site"
FT
XX
XX US2003022334-A1.
XX
XX 30-JAN-2003.
XX
XX 04-FEB-2002; 2002US-00067668.
XX
XX 02-FEB-2001; 2001US-0266140P.
XX (GLUC/) GLUCKSMANN M A.
XX
XX Glucksmann MA;
XX
XX WPI; 2003-447427/42.
XX N-PSDB; ACA61907.
XX
XX New 33312, 33303 or 32579 nucleic acid molecule encoding cytochrome P450 polypeptides, useful for diagnosing and treating cancer, inflammation, vascular disorders, CNS disorders or metabolic disorders, and in pharmacogenomics.
XX
XX Claim 5; Page 48; 76pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule comprising any of the 6 nucleic acid sequences, given in the specification which encodes a polypeptide having any of the 3 amino acid sequences, given in the specification. The nucleic acid molecule, protein and antibody are useful in screening assays and predictive medicine, such as in diagnostic assays, prognostic assays, monitoring clinical trials and pharmacogenetics. The composition and method are used in diagnosing and treating 33312, 33303 or 32579 cytochrome P450-related disorders, such as cellular proliferative and/or differentiative disorders (e.g. cancer), disorders associated with undesirable or deficient cell adhesion, motility or migration, inflammatory or autoimmune disorders (e.g. rheumatoid arthritis, diabetes, multiple sclerosis, systemic lupus erythematosus, conjunctivitis, allergies, asthma, Grave's disease or graft-versus-host disease), vascular disorders, central nervous system (CNS) disorders (e.g. Alzheimer's disease, anxiety disorders, amnesia, depression, migraine or Parkinson's disease), cell-signalling-associated disorders, metabolism-associated disorders (e.g. obesity, cachexia or anorexia), steroid-associated disorders and fatty acid-associated disorders. The present sequence represents the amino acid sequence of human cytochrome P450, 33303
XX
SQ Sequence 504 AA;
Query Match 100.0%; Score 2615; DB 6; Length 504;
Best Local Similarity 100.0%; Pred. No. 1.8e-258;
Matches 504; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MEATGTWALLLALLLLLLLTLALSGTRARGHLPQPTPLPLGNLLQLRPGALYSGLMRL 60
Db 1 MEATGTWALLLALLLLLLLTLALSGTRARGHLPQPTPLPLGNLLQLRPGALYSGLMRL 60
QY 61 SKKYGVPFTYLGPRPWWLVGQEAVERALGGQAEFFSGRTGVAMLEGTDFGHGVFFSN 120
Db 61 SKKYGVPFTYLGPRPWWLVGQEAVERALGGQAEFFSGRTGVAMLEGTDFGHGVFFSN 120
```





PR 06-JAN-2000; 2000WO-US000277.  
PR 06-JAN-2000; 2000WO-US000376.  
PR 11-FEB-2000; 2000WO-US003565.  
PR 18-FEB-2000; 2000WO-US004341.  
PR 18-FEB-2000; 2000WO-US004342.  
PR 22-FEB-2000; 2000WO-US004414.  
PR 24-FEB-2000; 2000WO-US004914.  
PR 24-FEB-2000; 2000WO-US005004.  
PR 01-MAR-2000; 2000WO-US005601.  
PR 02-MAR-2000; 2000WO-US005746.  
XX  
PA (GETH ) GENENTECH INC.  
XX  
XX Baker KP, Beresini M, Deforge L, Desnoyers L, Filvaroff E, Gao W;  
PI Gerritsen ME, Goddard A, Godowski FJ, Gurney AL, Sherwood S;  
PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;  
XX  
XX WPI; 2003-352836/33.  
DR N-P5DB; ACA67025.  
XX  
XX New isolated PRO polypeptide useful for treating diabetes, rheumatoid  
PT arthritis, sports injuries, obesity, hearing loss in mammals, stroke, or  
PT heart attack.  
XX  
XX Claim 12; Fig 64; 643pp; English.  
XX  
XX The present invention relates to the isolation of novel human PRO  
CC polypeptides, and the polynucleotide sequences encoding them. The PRO  
CC polypeptides are secreted and transmembrane proteins. The PRO  
CC polypeptides and polynucleotides are useful for preparing a medicament  
CC useful in the treatment of diabetes, bone and/or cartilage disorders  
CC (e.g. rheumatoid arthritis, sports injuries, osteoarthritis), obesity,  
CC hyper- or hypo-insulinaemia, hearing loss, and coagulation disorders  
CC assays for PRO, by detecting its expression in specific cells, tissues or  
CC serum, and for affinity purification of PRO from recombinant cell culture  
CC or natural sources. ABU8670-ABU81144 represent the human PRO  
CC polypeptides of the invention. Note: The sequence data for this patent  
CC was obtained in electronic format directly from the USPTO web site at  
CC seqdata.uspto.gov/psipsdIDEntry.html  
XX  
XX Sequence 504 AA;  
SQ  
Query Match 100.0%; Score 2615; DB 6; Length 504;  
Best Local Similarity 100.0%; Pred. No. 1.8e-258;  
Matches 504; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MEATGTWALLALALLLLTALSCTRARGHLPPTPLPLGMLQLRPGALYSGLMRL 60  
DB 1 MEATGTWALLALALLLLTALSCTRARGHLPPTPLPLGMLQLRPGALYSGLMRL 60  
QY 61 SKKYGPVFTIYLGFWPVRVVLVGOEAVREALGQOAEFEFGRTVAMLEGTFDGHGVFFSN 120  
DB 61 SKKYGPVFTIYLGFWPVRVVLVGOEAVREALGQOAEFEFGRTVAMLEGTFDGHGVFFSN 120  
QY 121 GERWRQLRKTMTALRDLGKGECEGLIOAEARCLVETFOGTGEPDPSLLLAQATSN 180  
DB 121 GERWRQLRKTMTALRDLGKGECEGLIOAEARCLVETFOGTGEPDPSLLLAQATSN 180  
QY 181 VCSLLFGLRFSYEDKEFOAVVRAAGTLLGVSSQGGQTYEMFSWFLRPLPGPHKQLLHH 240  
DB 181 VCSLLFGLRFSYEDKEFOAVVRAAGTLLGVSSQGGQTYEMFSWFLRPLPGPHKQLLHH 240  
QY 241 VSTLAFTVRQVQHQNLNDASGPARDLVDAFLKMAQEQNPGTFTNKNMLMTVYLL 300  
DB 241 VSTLAFTVRQVQHQNLNDASGPARDLVDAFLKMAQEQNPGTFTNKNMLMTVYLL 300  
QY 301 FAGTMTVSTTVGTYLLLMKYPHVQKWRBELNELGAGQAPSLGDRTRLPYTDVALLHEA 360  
DB 301 FAGTMTVSTTVGTYLLLMKYPHVQKWRBELNELGAGQAPSLGDRTRLPYTDVALLHEA 360  
QY 361 QRLALVPMGIPRTIMRTTRFRGYTLPGQTEVFFLLGSILHDPNIFKHPEEFNDRFLDA 420

Db 361 QRLALVPMGIPRTIMRTTRFRGYTLPGQTEVFFLLGSILHDPNIFKHPEEFNDRFLDA 420  
QY 421 DGRERKHEAFLPFSIGKRVCLGEGAKAELELFTTTIIQAFLSLESPCPDPTLSLKPTVSG 480  
Db 421 DGRFRKHEAFLPFSIGKRVCLGEGAKAELELFTTTIIQAFLSLESPCPDPTLSLKPTVSG 480  
QY 481 LFNIPPFAFQLOVRPTDLHSTTTQTR 504  
Db 481 LFNIPPFAFQLOVRPTDLHSTTTQTR 504  
RESULT 7  
ABU66601  
ID ABU66601 standard; protein; 504 AA.  
XX AC ABU66601;  
XX DT 23-MAY-2003 (first entry)  
XX DE Human PRO polypeptide #32.  
XX KW Human; PRO polypeptide; secreted and transmembrane protein;  
KW tumour necrosis factor-alpha; TNF-alpha; blood; proliferation;  
KW differentiation; chondrocyte; tumour; genetic disorder; cytostatic.  
XX OS Homo sapiens.  
XX PN US2003036180-A1.  
XX PD 20-FEB-2003.  
XX PF 09-MAY-2002; 2002US-00143114.  
XX PR 31-MAR-1997; 97WO-US0005230.  
PR 12-JUN-1998; 98WO-US012456.  
PR 14-JUL-1998; 98WO-US014552.  
PR 28-AUG-1998; 98WO-US017888.  
PR 10-SEP-1998; 98WO-US018824.  
PR 14-SEP-1998; 98WO-US019093.  
PR 14-SEP-1998; 98WO-US019094.  
PR 14-SEP-1998; 98WO-US019177.  
PR 16-SEP-1998; 98WO-US019330.  
PR 17-SEP-1998; 98WO-US019437.  
PR 07-OCT-1998; 98WO-US021141.  
PR 29-OCT-1998; 98WO-US022991.  
PR 29-OCT-1998; 98WO-US022992.  
PR 20-NOV-1998; 98WO-US024855.  
PR 01-DEC-1998; 98WO-US025108.  
PR 05-JAN-1999; 99WO-US000106.  
PR 08-MAR-1999; 99WO-US005028.  
PR 10-MAR-1999; 99WO-US005190.  
PR 20-APR-1999; 99WO-US008615.  
PR 14-MAY-1999; 99WO-US010733.  
PR 01-SEP-1999; 99WO-US012252.  
PR 01-SEP-1999; 99WO-US020111.  
PR 08-SEP-1999; 99WO-US020594.  
PR 13-SEP-1999; 99WO-US020944.  
PR 15-SEP-1999; 99WO-US021090.  
PR 15-SEP-1999; 99WO-US021547.  
PR 05-OCT-1999; 99WO-US023089.  
PR 29-NOV-1999; 99WO-US028214.  
PR 30-NOV-1999; 99WO-US028313.  
PR 30-NOV-1999; 99WO-US028409.  
PR 01-DEC-1999; 99WO-US028301.  
PR 01-DEC-1999; 99WO-US028634.  
PR 02-DEC-1999; 99WO-US028564.  
PR 02-DEC-1999; 99WO-US028565.  
PR 16-DEC-1999; 99WO-US030095.  
PR 20-DEC-1999; 99WO-US030911.  
PR 20-DEC-1999; 99WO-US030999.  
PR 22-DEC-1999; 99WO-US030720.  
PR 30-DEC-1999; 99WO-US031243.

PR 30-DEC-1999; 99WO-US031274.  
 PR 05-JAN-2000; 2000WO-US000219.  
 PR 06-JAN-2000; 2000WO-US000217.  
 PR 06-JAN-2000; 2000WO-US000376.  
 PR 11-FEB-2000; 2000WO-US0003565.  
 PR 18-FEB-2000; 2000WO-US004341.  
 PR 18-FEB-2000; 2000WO-US004342.  
 PR 22-FEB-2000; 2000WO-US004414.  
 PR 24-FEB-2000; 2000WO-US004914.  
 PR 24-FEB-2000; 2000WO-US005004.  
 PR 01-MAR-2000; 2000WO-US005601.  
 PR 02-MAR-2000; 2000WO-US005746.  
 PR 02-MAR-2000; 2000WO-US005841.  
 PR 10-MAR-2000; 2000WO-US006319.  
 PR 15-MAR-2000; 2000WO-US006884.  
 PR 20-MAR-2000; 2000WO-US007377.  
 PR 21-MAR-2000; 2000WO-US007532.  
 PR 30-MAR-2000; 2000WO-US008439.  
 PR 17-MAY-2000; 2000WO-US013705.  
 PR 22-MAY-2000; 2000WO-US014042.  
 PR 30-MAY-2000; 2000WO-US014941.  
 PR 02-JUN-2000; 2000WO-US015264.  
 PR 28-JUL-2000; 2000WO-US020710.  
 PR 11-AUG-2000; 2000WO-US022031.  
 PR 23-AUG-2000; 2000WO-US023522.  
 PR 24-AUG-2000; 2000WO-US023328.  
 PR 08-NOV-2000; 2000WO-US030952.  
 PR 10-NOV-2000; 2000WO-US030873.  
 PR 01-DEC-2000; 2000WO-US032678.  
 PR 20-DEC-2000; 2000US-00747259.  
 PR 20-DEC-2000; 2000WO-US034956.  
 PR 28-FEB-2001; 2001US-00796498.  
 PR 28-FEB-2001; 2001WO-US006520.  
 PR 01-MAR-2001; 2001WO-US006666.  
 PR 09-MAR-2001; 2001US-00802706.  
 PR 14-MAR-2001; 2001US-00806889.  
 PR 22-MAR-2001; 2001US-00816744.  
 PR 03-APR-2001; 2001US-00828366.  
 PR 10-MAY-2001; 2001US-00854208.  
 PR 10-MAY-2001; 2001US-00854280.  
 PR 18-MAY-2001; 2001US-00860216.  
 PR 25-MAY-2001; 2001US-00866028.  
 PR 25-MAY-2001; 2001US-00866034.  
 PR 25-MAY-2001; 2001WO-US017092.  
 PR 01-JUN-2001; 2001US-00872035.  
 PR 01-JUN-2001; 2001WO-US017800.  
 PR 05-JUN-2001; 2001US-00874503.  
 PR 14-JUN-2001; 2001US-00882636.  
 PR 19-JUN-2001; 2001US-00886342.  
 PR 20-JUN-2001; 2001WO-US019692.  
 PR 21-JUN-2001; 2001US-00887879.  
 PR 22-JUN-2001; 2001WO-US020116.  
 PR 29-JUN-2001; 2001WO-US021066.  
 PR 09-JUL-2001; 2001WO-US021735.  
 PR 18-JUL-2001; 2001US-00908827.  
 PR 06-AUG-2001; 2001US-00924419.  
 PR 09-AUG-2001; 2001US-00927796.  
 PR 16-AUG-2001; 2001US-00931836.  
 PR 19-DEC-2001; 2001US-00028072.  
 (GETH ) GENENTECH INC.

PA Baker KP, Beresini M, Deforge L, Desnoyers L, Filvaroff E, Gao W;  
 PI Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;  
 PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;  
 DR WPI: 2003-332040/31.  
 DR N-PSDB; ACA03634.  
 XX  
 PT New secreted and transmembrane PRO nucleic acids, useful for gene  
 PT therapy, in chromosome and gene mapping, as chromosome markers, in tissue  
 PT typing, and in chromosome identification.  
 XX

PS Claim 12; Fig 64; 660pp; English.  
 XX The present invention relates to the isolation of novel human PRO  
 CC polypeptides, and the polynucleotide sequences encoding them. The PRO  
 CC polypeptides are secreted and transmembrane proteins. The PRO  
 CC polypeptides are useful for detecting other PRO polypeptides, for linking  
 CC bioactive molecules to cells expressing PRO polypeptides, for modulating  
 CC biological activities of cells expressing PRO polypeptides, and for  
 CC identifying agonists or antagonists. The PRO polypeptides are useful for  
 CC for stimulating the release of tumour necrosis factor (TNF)-alpha from  
 CC human blood, for stimulating the proliferation or differentiation of  
 CC chondrocytes, and detecting the presence of tumours. The polynucleotide  
 CC sequences encoding PRO polypeptides are useful as hybridisation probes,  
 CC in chromosome and gene mapping, in the generation of antisense RNA and  
 CC DNA, in the preparation of PRO polypeptides, for generating transgenic  
 CC animals or knockout animals, for the genetic analysis of individuals with  
 CC genetic disorders, and in gene therapy. ABU6570-ABU6684 represent the  
 CC human PRO polypeptides of the invention. Note: The sequence data for this  
 CC patent was obtained in electronic format directly from the USPTO web site  
 CC at seqdata.uspto.gov/paipedEntry.html  
 XX  
 SQ Sequence 504 AA;  
 Query Match 100.0%; Score 2615; DB 6; Length 504;  
 Best Local Similarity 100.0%; Pred. No. 1.8e-258;  
 Matches 504; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 MEATGTWALLALALLLLTLLSGLTRARGHLPPGTPPLPLGGLLQIRPGALYGLMRL 60  
 Db 1 MEATGTWALLALALLLLTLLSGLTRARGHLPPGTPPLPLGGLLQIRPGALYGLMRL 60  
 Qy 61 SKYGPVFTIYLGPRPVVVLVGQEAVERALGGQAEFSGRGTVAMLEGTFDGHGVFFSN 120  
 Db 61 SKYGPVFTIYLGPRPVVVLVGQEAVERALGGQAEFSGRGTVAMLEGTFDGHGVFFSN 120  
 Qy 121 GERWQLRKFTMLALRDGLMGKREGEELIOAEARCLVETFGTEGRFPDPSLLAQATSN 180  
 Db 121 GERWQLRKFTMLALRDGLMGKREGEELIOAEARCLVETFGTEGRFPDPSLLAQATSN 180  
 Qy 181 VVCSLLFGLRFSYEDKEFOAVVRAAGTLLGVSSQGGQTYEMFSWFLRPLPGHKQLLHH 240  
 Db 181 VVCSLLFGLRFSYEDKEFOAVVRAAGTLLGVSSQGGQTYEMFSWFLRPLPGHKQLLHH 240  
 Qy 241 VSTLAFTVRQVQOQHGNLDASGPARDLVDAFLKWAQEEQNPGTEFTKNMLMTVIYLL 300  
 Db 241 VSTLAFTVRQVQOQHGNLDASGPARDLVDAFLKWAQEEQNPGTEFTKNMLMTVIYLL 300  
 Qy 301 FAGTMTVSTTVGYTLLMLMKYHVQKVRRELNRRELGAQAPSLGDRTRLPYTDVILHEA 360  
 Db 301 FAGTMTVSTTVGYTLLMLMKYHVQKVRRELNRRELGAQAPSLGDRTRLPYTDVILHEA 360  
 Qy 361 QRLLLVPMGIPRTLMRTTRFRGYTLPGQTEVPFLGSLTLDHPNIFKHPEENPDRFLDA 420  
 Db 361 QRLLLVPMGIPRTLMRTTRFRGYTLPGQTEVPFLGSLTLDHPNIFKHPEENPDRFLDA 420  
 Qy 421 DGRFRKHEAFLPFLSKGRVCLGEGAKAELFFFTTILQAFSLSEPCPDDTSLKPTVSG 480  
 Db 421 DGRFRKHEAFLPFLSKGRVCLGEGAKAELFFFTTILQAFSLSEPCPDDTSLKPTVSG 480  
 Qy 481 LFNIPPAFQOLVRPTDLHSTTTQTR 504  
 Db 481 LFNIPPAFQOLVRPTDLHSTTTQTR 504  
 RESULT 8  
 ABUS9682  
 ID ABUS9682 standard; protein; 504 AA.  
 XX  
 AC ABUS9682;  
 XX  
 DT 13-MAY-2003 (first entry)  
 XX  
 DE Novel secreted and transmembrane protein PRO1906.

XX Human; PRO; hypertrophy of neonatal heart; angiogenesis; wound healing;  
KW cardiac insufficiency disorder; cancer; tumour; immune response;  
KW adrenal cortical capillary endothelial growth; c-fos induction;  
KW vascular endothelial growth factor inhibition; VEGF inhibition;  
KW endothelial cell growth inhibitor; T-lymphocytes stimulation;  
KW retinal neurons cell survival; rod photoreceptor cell survival;  
KW retinal disorder; retinitis pigmentosa; kidney disease;  
KW mammalian kidney mesangial cell proliferation; Berger disease;  
KW dermatitis; herpeticiformis; Crohn's disease; chondrocyte proliferation;  
KW chondrocyte redifferentiation; sports injury; arthritis.  
XX  
OS Homo sapiens.  
XX  
PN US2003017563-A1.  
XX  
PD 23-JAN-2003.  
XX  
PF 07-MAY-2002; 2002US-00140808.  
XX  
XX 31-MAR-1997; 97WO-US005230.  
PR 12-JUN-1998; 98WO-US012456.  
PR 14-JUL-1998; 98WO-US014552.  
PR 28-AUG-1998; 98WO-US017888.  
PR 10-SEP-1998; 98WO-US018824.  
PR 14-SEP-1998; 98WO-US019093.  
PR 14-SEP-1998; 98WO-US019094.  
PR 14-SEP-1998; 98WO-US019177.  
PR 16-SEP-1998; 98WO-US019330.  
PR 17-SEP-1998; 98WO-US019437.  
PR 07-OCT-1998; 98WO-US021144.  
PR 29-OCT-1998; 98WO-US022991.  
PR 29-OCT-1998; 98WO-US022992.  
PR 20-NOV-1998; 98WO-US024855.  
PR 01-DEC-1998; 98WO-US025108.  
PR 05-JAN-1999; 98WO-US000106.  
PR 08-MAR-1999; 98WO-US005028.  
PR 10-MAR-1999; 98WO-US005190.  
PR 20-APR-1999; 98WO-US008615.  
PR 14-MAY-1999; 98WO-US010733.  
PR 02-JUN-1999; 98WO-US012252.  
PR 01-SEP-1999; 98WO-US020111.  
PR 08-SEP-1999; 98WO-US020594.  
PR 13-SEP-1999; 98WO-US020594.  
PR 15-SEP-1999; 98WO-US021090.  
PR 15-SEP-1999; 98WO-US021547.  
PR 05-OCT-1999; 98WO-US023089.  
PR 30-NOV-1999; 98WO-US028214.  
PR 30-NOV-1999; 98WO-US028313.  
PR 30-NOV-1999; 98WO-US028409.  
PR 01-DEC-1999; 98WO-US028301.  
PR 01-DEC-1999; 98WO-US028634.  
PR 02-DEC-1999; 98WO-US028551.  
PR 02-DEC-1999; 98WO-US028564.  
PR 02-DEC-1999; 98WO-US028565.  
PR 16-DEC-1999; 98WO-US030095.  
PR 20-DEC-1999; 98WO-US030911.  
PR 22-DEC-1999; 98WO-US030999.  
PR 22-DEC-1999; 98WO-US030720.  
PR 30-DEC-1999; 98WO-US031243.  
PR 30-DEC-1999; 98WO-US031274.  
PR 03-JAN-2000; 2000WO-US000219.  
PR 06-JAN-2000; 2000WO-US000277.  
PR 06-JAN-2000; 2000WO-US000376.  
PR 11-FEB-2000; 2000WO-US003565.  
PR 18-FEB-2000; 2000WO-US004341.  
PR 18-FEB-2000; 2000WO-US004342.  
PR 22-FEB-2000; 2000WO-US004414.  
PR 24-FEB-2000; 2000WO-US004514.  
PR 24-FEB-2000; 2000WO-US005004.  
PR 01-MAR-2000; 2000WO-US005601.  
PR 02-MAR-2000; 2000WO-US005746.  
PR 02-MAR-2000; 2000WO-US005841.  
PR 10-MAR-2000; 2000WO-US006319.  
PR 15-MAR-2000; 2000WO-US006884.  
PR 20-MAR-2000; 2000WO-US007377.  
PR 21-MAR-2000; 2000WO-US007532.  
PR 30-MAR-2000; 2000WO-US008439.  
PR 17-MAY-2000; 2000WO-US013705.  
PR 22-MAY-2000; 2000WO-US014042.  
PR 30-MAY-2000; 2000WO-US014941.  
PR 02-JUN-2000; 2000WO-US015264.  
PR 28-JUL-2000; 2000WO-US020710.  
PR 11-AUG-2000; 2000WO-US022031.  
PR 23-AUG-2000; 2000WO-US023522.  
PR 24-AUG-2000; 2000WO-US023328.  
PR 08-NOV-2000; 2000WO-US030952.  
PR 10-NOV-2000; 2000WO-US030873.  
PR 01-DEC-2000; 2000WO-US032678.  
PR 20-DEC-2000; 2000WO-US034956.  
PR 20-DEC-2000; 2000WO-US034956.  
PR 28-FEB-2001; 2001US-00796498.  
PR 28-FEB-2001; 2001WO-US006520.  
PR 01-MAR-2001; 2001WO-US006666.  
PR 09-MAR-2001; 2001US-00802706.  
PR 14-MAR-2001; 2001US-00808689.  
PR 22-MAR-2001; 2001US-00816744.  
PR 05-APR-2001; 2001US-00828366.  
PR 10-MAY-2001; 2001US-00854208.  
PR 10-MAY-2001; 2001US-00854280.  
PR 18-MAY-2001; 2001US-00860216.  
PR 25-MAY-2001; 2001US-00866028.  
PR 25-MAY-2001; 2001US-00866034.  
PR 25-MAY-2001; 2001WO-US017092.  
PR 01-JUN-2001; 2001US-00872035.  
PR 01-JUN-2001; 2001WO-US017800.  
PR 05-JUN-2001; 2001US-00874503.  
PR 14-JUN-2001; 2001US-00882636.  
PR 19-JUN-2001; 2001US-00886342.  
PR 20-JUN-2001; 2001WO-US019692.  
PR 21-JUN-2001; 2001US-00887879.  
PR 22-JUN-2001; 2001WO-US020116.  
PR 29-JUN-2001; 2001WO-US021066.  
PR 09-JUL-2001; 2001WO-US021735.  
PR 18-JUL-2001; 2001US-00908827.  
PR 06-AUG-2001; 2001US-00924419.  
PR 09-AUG-2001; 2001US-00927796.  
PR 16-AUG-2001; 2001US-00931836.  
PR 19-DEC-2001; 2001US-00028072.  
XX  
XX (GETH ) GENENTECH INC.  
PA Baker KP, Beresini M, Deforge L, Desnoyers L, Filvaroff E, Gao W;  
PI Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;  
PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;  
XX  
XX WPI; 2003-148238/14.  
DR N-PSDB; ABX89172.  
XX  
XX Two hundred and seventy five nucleic acids encoding PRO polypeptides,  
PT useful for treating pericyte-associated tumors, diabetes and various bone  
PT and/or cartilage disorders, e.g. arthritis.  
XX  
XX Claim 12; Fig 64; 65pp; English.  
XX  
XX The invention describes an isolated human PRO polypeptide. The PRO  
CC polypeptides are useful in detecting PRO polypeptides in a sample, in  
CC linking a bioactive molecule to a cell expressing a PRO polypeptide, and  
CC in modulating at least one biological activity of a cell expressing a PRO  
CC polypeptide. PRO1312 stimulates hypertrophy of neonatal heart and is thus  
CC useful for treating cardiac insufficiency disorders. PRO1154 and PRO1186  
CC stimulate adrenal cortical capillary endothelial growth, and PRO536,  
CC PRO943, PRO826, PRO1068 or PRO535, PRO826, PRO819, PRO1126,  
CC PRO1360 and PRO1387 induce c-fos in endothelial cells, and are thus  
CC useful for treating conditions or disorders where angiogenesis would be  
CC beneficial, e.g. wound healing and antagonist of this polypeptide are

CC useful for treating cancerous tumours. PRO812 inhibits vascular  
CC endothelial growth factor (VEGF) stimulated proliferation of endothelial  
CC cells and is thus useful for inhibiting endothelial cell growth in  
CC mammals which would be beneficial in inhibiting tumour growth. PRO826,  
CC PRO1068, PRO1184, PRO1346 and PRO1375 stimulate proliferation of  
CC stimulated T-lymphocytes and are therapeutically useful for enhancing  
CC immune response. PRO828, PRO826, PRO1068 or PRO132 enhance survival of  
CC retinal neurons cells (PRO1132 is also enhances survival/proliferation of  
CC rod photoreceptor cells) and therefore are useful for treating retinal  
CC disorders of injuries, e.g. retinitis pigmentosa, AMD. PRO819, PRO813  
CC and PRO1066 induce proliferation of mammalian kidney mesangial cells,  
CC and therefore are useful for treating kidney disorders associated with  
CC decreased mesangial cell function such as Berger disease or other  
CC nephropathies associated with dermatitis, herpetiformis or Crohn's  
CC disease. PRO1310, PRO844, PRO1312, PRO1192 and PRO1387 induce the  
CC proliferation and/or redifferentiation of chondrocytes in culture and are  
CC thus useful for treating sports injuries, and arthritis. This is the  
CC amino acid sequence of a novel human PRO protein  
XX Sequence 504 AA;

Query Match 100.0%; Score 2615; DB 6; Length 504;  
Best Local Similarity 100.0%; Pred. No. 1.8e-258; Indels 0; Gaps 0;  
Matches 504; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MEATGTWALLALALLLLTALSGTRAGHLPFGTPLPLGNNLQLRPGALYGLMRL 60  
Db 1 MEATGTWALLALALLLLTALSGTRAGHLPFGTPLPLGNNLQLRPGALYGLMRL 60  
Qy 61 SKKYGVPFTIYLGWPRVWVVGQAVREALGQAEFSGRTVAMLEGTFDGHGVFFSN 120  
Db 61 SKKYGVPFTIYLGWPRVWVVGQAVREALGQAEFSGRTVAMLEGTFDGHGVFFSN 120  
Qy 121 GERWQLRKFTMLALRDLGMKRGEEELIQAEARCLVETFGTEGRPDPSLLLAQATSN 180  
Db 121 GERWQLRKFTMLALRDLGMKRGEEELIQAEARCLVETFGTEGRPDPSLLLAQATSN 180  
Qy 181 VVCSLLFGLRFSYEDKEFQAVVRAAGTLLGVSSQGGQTYEMFSWFLRPLPGPHKQLLHH 240  
Db 181 VVCSLLFGLRFSYEDKEFQAVVRAAGTLLGVSSQGGQTYEMFSWFLRPLPGPHKQLLHH 240  
Qy 241 VSTLAFTVROVQHQNLDSAGPARDLVDAFLKMAQEEQNPGTEFTNKNMLMTVIYLL 300  
Db 241 VSTLAFTVROVQHQNLDSAGPARDLVDAFLKMAQEEQNPGTEFTNKNMLMTVIYLL 300  
Qy 301 FACTMTVSTTVGYTLMLLMKYPHVQKWVREELNRELGAQAPSLGDRTRLPYTDVILHEA 360  
Db 301 FACTMTVSTTVGYTLMLLMKYPHVQKWVREELNRELGAQAPSLGDRTRLPYTDVILHEA 360  
Qy 361 ORLLALVPMGIPRTLMTTRFRGYTLPGQTEVFPLLSILHDPNIPKHPEEFPNDRFLDA 420  
Db 361 ORLLALVPMGIPRTLMTTRFRGYTLPGQTEVFPLLSILHDPNIPKHPEEFPNDRFLDA 420  
Qy 421 DGRFRKHEAFPLPSLGRVCLGSLKAEALFLPFTTILQAFSLSPCPDPTLSLKTPTVSG 480  
Db 421 DGRFRKHEAFPLPSLGRVCLGSLKAEALFLPFTTILQAFSLSPCPDPTLSLKTPTVSG 480  
Qy 481 LFNIPPAFQLQVRPTDLHSTTQTR 504  
Db 481 LFNIPPAFQLQVRPTDLHSTTQTR 504

RESULT 9  
ABO24872  
ID ABO24872 standard; protein; 504 AA.  
XX AC ABO24872;  
XX AC ABO24872;  
DT 05-SEP-2003 (first entry)  
DE Human secreted/transmembrane protein (PRO) #32.  
XX Human; PRO; secreted protein; transmembrane protein; tumour; cytostatic;

gene therapy; tumour necrosis factor-alpha; TNF-alpha; blood;  
KW proteoglycan; cartilage; cytokine; peripheral blood mononuclear cell;  
KW PBMC; Glucose uptake; FFA; skeletal muscle cell; adipocyte cell;  
KW chondrocyte cell proliferation; chondrocyte cell differentiation;  
KW pericyte cell; inner ear utricular supporting cell; T-lymphocyte cell;  
KW endothelial cell; A-peptide; factor VIIA.  
XX Homo sapiens.  
XX US2003036179-A1.  
XX 20-FEB-2003.  
PF 10-MAY-2002; 2002US-00142431.  
XX 31-MAR-1997; 97WO-US005230.  
PR 12-JUN-1998; 98WO-US012456.  
PR 14-JUL-1998; 98WO-US014552.  
PR 28-AUG-1998; 98WO-US017888.  
PR 10-SEP-1998; 98WO-US018824.  
PR 14-SEP-1998; 98WO-US019093.  
PR 14-SEP-1998; 98WO-US019094.  
PR 14-SEP-1998; 98WO-US019177.  
PR 16-SEP-1998; 98WO-US019330.  
PR 17-SEP-1998; 98WO-US019437.  
PR 07-OCT-1998; 98WO-US021141.  
PR 29-OCT-1998; 98WO-US022991.  
PR 29-OCT-1998; 98WO-US022992.  
PR 20-NOV-1998; 98WO-US024855.  
PR 01-DEC-1998; 98WO-US025108.  
PR 05-JAN-1999; 99WO-US000106.  
PR 08-MAR-1999; 99WO-US005190.  
PR 10-MAR-1999; 99WO-US008615.  
PR 20-APR-1999; 99WO-US008615.  
PR 14-MAY-1999; 99WO-US010733.  
PR 02-JUN-1999; 99WO-US012252.  
PR 01-SEP-1999; 99WO-US020111.  
PR 08-SEP-1999; 99WO-US020594.  
PR 13-SEP-1999; 99WO-US020944.  
PR 15-SEP-1999; 99WO-US021090.  
PR 15-SEP-1999; 99WO-US021547.  
PR 03-OCT-1999; 99WO-US023089.  
PR 23-NOV-1999; 99WO-US028214.  
PR 30-NOV-1999; 99WO-US028313.  
PR 30-NOV-1999; 99WO-US028409.  
PR 01-DEC-1999; 99WO-US028301.  
PR 01-DEC-1999; 99WO-US028634.  
PR 02-DEC-1999; 99WO-US028551.  
PR 02-DEC-1999; 99WO-US028564.  
PR 02-DEC-1999; 99WO-US028565.  
PR 16-DEC-1999; 99WO-US030095.  
PR 16-DEC-1999; 99WO-US030911.  
PR 22-DEC-1999; 99WO-US030999.  
PR 22-DEC-1999; 99WO-US030720.  
PR 30-DEC-1999; 99WO-US031243.  
PR 30-DEC-1999; 99WO-US031274.  
PR 05-JAN-2000; 2000WO-US000219.  
PR 06-JAN-2000; 2000WO-US000277.  
PR 06-JAN-2000; 2000WO-US000376.  
PR 11-FEB-2000; 2000WO-US003565.  
PR 18-FEB-2000; 2000WO-US004341.  
PR 18-FEB-2000; 2000WO-US004342.  
PR 22-FEB-2000; 2000WO-US004414.  
PR 24-FEB-2000; 2000WO-US004914.  
PR 24-FEB-2000; 2000WO-US005004.  
PR 01-MAR-2000; 2000WO-US005601.  
PR 02-MAR-2000; 2000WO-US005746.  
PR 02-MAR-2000; 2000WO-US005841.  
PR 10-MAR-2000; 2000WO-US006319.  
PR 15-MAR-2000; 2000WO-US006884.  
PR 20-MAR-2000; 2000WO-US007377.  
PR 21-MAR-2000; 2000WO-US007532.  
PR 30-MAR-2000; 2000WO-US008439.

PRO polypeptide, modulating at least one biological activity of a cell expressing a PRO polypeptide, stimulating the release of tumour necrosis factor- $\alpha$  (TNF- $\alpha$ ) from human blood, (or proteoglycans from cartilage or cytokine from peripheral blood mononuclear cells (PBMC)), modulating the uptake of glucose or FFA by skeletal muscle cells or adipocyte cells, stimulating the proliferation or gene expression in pericyte chondrocyte cells (or proliferation of or gene expression in pericyte cells), stimulating the proliferation of inner ear utricular supporting cells, stimulating the proliferation of or endothelial cells), inhibiting the cells (or of T-lymphocyte cells, or of endothelial cells), or differentiation of adipocyte cells (or of A-peptide to factor VIIa, or differentiation of an cells, detecting the presence of a tumour in a mammal and an oligonucleotide probe derived from any of the nucleotide sequences given in the specification. The polynucleotide is useful in molecular biology, in the uses as hybridisation probes, in chromosome and gene mapping, including uses as hybridisation probes, and in gene therapy. The polynucleotide may also be used in preparing PRO polypeptides by recombinant techniques, and in generating either transgenic animals or knock-out animals which, in turn, are useful in the development and screening of therapeutically useful reagents. The PRO polypeptide or antibody is used in preparing a medicament for treating a condition responsive to the polypeptide or antibody, such as tumours, and in various diagnostic assays. The present sequence represents a PRO polypeptide

[illegible]

Qy	1	MEATG	WALLALALALLLLTALSGT	REARGHLP	PGPTPLPL	LLGNLQ	LRPGALYSGLMRL	60	
Db	1	MEATG	WALLALALALLLLTALSGT	REARGHLP	PGPTPLPL	LLGNLQ	LRPGALYSGLMRL	60	
Qy	61	SKXG	PVFTIYLG	PWRVWVLVG	QEA	VREALG	QQAEEFSGRGTVAMLEGTDHG	VPFSN 120	
Db	61	SKXG	PVFTIYLG	PWRVWVLVG	QEA	VREALG	QQAEEFSGRGTVAMLEGTDHG	VPFSN 120	
Qy	121	GERW	QRUKFTML	ALROLG	KGKREGEELI	QAEARCL	VETFOGTSGRFPDP	PSLLLAQAATSN 180	
Db	121	GERW	QRUKFTML	ALROLG	KGKREGEELI	QAEARCL	VETFOGTSGRFPDP	PSLLLAQAATSN 180	
Qy	181	VVCS	LLFGLRFSYEDKE	FQAV	VRAAGGTL	LGVSQGGQTYEMF	SWFRLRPL	PGPHKQLLHH 240	
Db	181	VVCS	LLFGLRFSYEDKE	FQAV	VRAAGGTL	LGVSQGGQTYEMF	SWFRLRPL	PGPHKQLLHH 240	
Qy	241	VSTL	AAFTVRVOQ	HOQNLD	DASGPARDI	VDAPFL	KQAQEEONPGTEFTN	KNMLMTVIYLL 300	
Db	241	VSTL	AAFTVRVOQ	HOQNLD	DASGPARDI	VDAPFL	KQAQEEONPGTEFTN	KNMLMTVIYLL 300	
Qy	301	FAGT	MTVSTTVGYITLL	LMKY	PHVQKWV	REELNREL	GAQAPSLGDRTRL	PYTDVAVLHEA 360	
Db	301	FAGT	MTVSTTVGYITLL	LMKY	PHVQKWV	REELNREL	GAQAPSLGDRTRL	PYTDVAVLHEA 360	
Qy	361	QRL	JALVPMGIPRTLM	TRFRGYTL	PGTEVFP	LLGSILHDP	PNIFKHP	EENPDRFLDA 420	
Db	361	QRL	JALVPMGIPRTLM	TRFRGYTL	PGTEVFP	LLGSILHDP	PNIFKHP	EENPDRFLDA 420	
Qy	421	DGR	FRKHEAF	PFSLGK	VCLGEG	LAKAEL	FFFTTIL	QAFSLES	PCPDDTUSLKPTVSG 480
Db	421	DGR	FRKHEAF	PFSLGK	VCLGEG	LAKAEL	FFFTTIL	QAFSLES	PCPDDTUSLKPTVSG 480
Qy	481	LFNI	PPAFOLQVR	PTDL	HS	TTQTR	504		
Db	481	LFNI	PPAFOLQVR	PTDL	HS	TTQTR	504		

RESULT 10  
ABU66877  
ID ABU66877 standard; protein: 504 AA.  
XX  
XX ABU66877;  
XX

17-MAY-2000; 2000WO-US013705.  
 22-MAY-2000; 2000WO-US014042.  
 30-MAY-2000; 2000WO-US014941.  
 02-JUN-2000; 2000WO-US015264.  
 28-JUL-2000; 2000WO-US020710.  
 11-AUG-2000; 2000WO-US022031.  
 23-AUG-2000; 2000WO-US023522.  
 24-AUG-2000; 2000WO-US023328.  
 08-NOV-2000; 2000WO-US030952.  
 10-NOV-2000; 2000WO-US030873.  
 01-DEC-2000; 2000WO-US032678.  
 20-DEC-2000; 2000US-00747259.  
 20-DEC-2000; 2000WO-US034956.  
 28-FEB-2001; 2001US-00796498.  
 28-FEB-2001; 2001WO-US006520.  
 01-MAR-2001; 2001WO-US006666.  
 09-MAR-2001; 2001US-00802706.  
 14-MAR-2001; 2001US-00808689.  
 22-MAR-2001; 2001US-00816744.  
 05-APR-2001; 2001US-00828366.  
 10-MAY-2001; 2001US-00854208.  
 10-MAY-2001; 2001US-00854280.  
 18-MAY-2001; 2001US-00860216.  
 25-MAY-2001; 2001US-00866028.  
 25-MAY-2001; 2001US-00866034.  
 25-MAY-2001; 2001WO-US017092.  
 01-JUN-2001; 2001US-00872035.  
 01-JUN-2001; 2001WO-US017800.  
 05-JUN-2001; 2001US-00874503.  
 14-JUN-2001; 2001US-00882636.  
 19-JUN-2001; 2001US-00886342.  
 20-JUN-2001; 2001WO-US019692.  
 21-JUN-2001; 2001US-00887879.  
 22-JUN-2001; 2001WO-US020116.  
 29-JUN-2001; 2001WO-US021066.  
 09-JUL-2001; 2001WO-US021735.  
 18-JUL-2001; 2001US-00908827.  
 06-AUG-2001; 2001US-00924419.  
 09-AUG-2001; 2001US-00927796.  
 16-AUG-2001; 2001US-00931836.  
 19-DEC-2001; 2001US-00028072.  
 XX  
 XX (GETH ) GENENTECH INC.  
 PA Baker KP, Beresini M, Deforge L, Desnoyers L, Filvaroff E, Gao W;  
 PI Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;  
 PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WJ, Zhang Z;  
 XX  
 XX WPI; 2003-466355/44.  
 DR N-PSDB; ACD41826.  
 DR  
 XX  
 XX New isolated nucleic acid encoding a PRO polypeptide, e.g. PRO1114 or  
 PT PRO4978, useful in molecular biology, chromosome and gene mapping, in  
 PT screening antisense RNA and DNA, and in gene therapy.  
 TT

27-MAY-2003 (first entry)  
Human secreted/transmembrane, PRO, protein SEQ ID 64.  
Human; secreted protein; transmembrane protein; PRO;  
inflammatory disease; organ failure; atherosclerosis; cardiac injury;  
infertility; birth defects; premature aging; AIDS; biosensor;  
acquired immunodeficiency syndrome; cancer; diabetic complication;  
bioreactor; tumour.  
Homo sapiens.  
US2003032155-A1.  
13-FEB-2003.  
03-MAY-2002; 2002US-00137865.  
31-MAR-1997; 97WO-US005230.  
12-JUN-1998; 98WO-US012456.  
14-JUL-1998; 98WO-US014552.  
28-AUG-1998; 98WO-US017888.  
10-SEP-1998; 98WO-US018824.  
14-SEP-1998; 98WO-US019093.  
14-SEP-1998; 98WO-US019094.  
14-SEP-1998; 98WO-US019177.  
16-SEP-1998; 98WO-US019330.  
17-SEP-1998; 98WO-US019437.  
07-OCT-1998; 98WO-US021141.  
29-OCT-1998; 98WO-US022991.  
29-OCT-1998; 98WO-US022992.  
20-NOV-1998; 98WO-US024855.  
01-DEC-1998; 98WO-US025108.  
05-JAN-1999; 99WO-US000106.  
08-MAR-1999; 99WO-US005028.  
10-MAR-1999; 99WO-US005190.  
20-APR-1999; 99WO-US008615.  
14-MAY-1999; 99WO-US010733.  
02-JUN-1999; 99WO-US012252.  
01-SEP-1999; 99WO-US020111.  
08-SEP-1999; 99WO-US020594.  
13-SEP-1999; 99WO-US020944.  
15-SEP-1999; 99WO-US021090.  
15-SEP-1999; 99WO-US021547.  
05-OCT-1999; 99WO-US023089.  
29-NOV-1999; 99WO-US028214.  
30-NOV-1999; 99WO-US028313.  
30-NOV-1999; 99WO-US028409.  
01-DEC-1999; 99WO-US028301.  
01-DEC-1999; 99WO-US028634.  
02-DEC-1999; 99WO-US028551.  
02-DEC-1999; 99WO-US028564.  
02-DEC-1999; 99WO-US028565.  
16-DEC-1999; 99WO-US030095.  
20-DEC-1999; 99WO-US030911.  
20-DEC-1999; 99WO-US030999.  
22-DEC-1999; 99WO-US030720.  
30-DEC-1999; 99WO-US031243.  
30-DEC-1999; 99WO-US031274.  
05-JAN-2000; 2000WO-US000219.  
06-JAN-2000; 2000WO-US000277.  
06-JAN-2000; 2000WO-US000376.  
11-FEB-2000; 2000WO-US003565.  
18-FEB-2000; 2000WO-US004341.  
18-FEB-2000; 2000WO-US004342.  
22-FEB-2000; 2000WO-US004414.  
24-FEB-2000; 2000WO-US004914.  
24-FEB-2000; 2000WO-US005004.  
01-MAR-2000; 2000WO-US005601.  
02-MAR-2000; 2000WO-US005746.  
02-MAR-2000; 2000WO-US005841.  
10-MAR-2000; 2000WO-US006319.  
15-MAR-2000; 2000WO-US006884.  
20-MAR-2000; 2000WO-US007377.  
21-MAR-2000; 2000WO-US007532.  
30-MAR-2000; 2000WO-US008439.  
17-MAY-2000; 2000WO-US013705.  
22-MAY-2000; 2000WO-US014042.  
30-MAY-2000; 2000WO-US014941.  
02-JUN-2000; 2000WO-US015264.  
28-JUL-2000; 2000WO-US020710.  
11-AUG-2000; 2000WO-US022031.  
23-AUG-2000; 2000WO-US023522.  
24-AUG-2000; 2000WO-US023328.  
08-NOV-2000; 2000WO-US030952.  
10-NOV-2000; 2000WO-US030873.  
01-DEC-2000; 2000WO-US032678.  
20-DEC-2000; 2000US-00747259.  
20-DEC-2000; 2000WO-US034956.  
28-FEB-2001; 2001US-00796498.  
28-FEB-2001; 2001WO-US006520.  
01-MAR-2001; 2001WO-US006666.  
09-MAR-2001; 2001US-00802706.  
14-MAR-2001; 2001US-00808689.  
22-MAR-2001; 2001US-00816744.  
05-APR-2001; 2001US-00828366.  
10-MAY-2001; 2001US-00854208.  
10-MAY-2001; 2001US-00854280.  
18-MAY-2001; 2001US-00860216.  
25-MAY-2001; 2001US-00866028.  
25-MAY-2001; 2001US-00866034.  
25-MAY-2001; 2001WO-US017092.  
01-JUN-2001; 2001US-00872035.  
01-JUN-2001; 2001WO-US017800.  
05-JUN-2001; 2001US-00874503.  
14-JUN-2001; 2001US-00882636.  
19-JUN-2001; 2001US-00886342.  
20-JUN-2001; 2001WO-US019692.  
21-JUN-2001; 2001US-00887879.  
22-JUN-2001; 2001WO-US020116.  
29-JUN-2001; 2001WO-US021066.  
09-JUL-2001; 2001WO-US021735.  
18-JUL-2001; 2001US-00908827.  
06-AUG-2001; 2001US-00924419.  
09-AUG-2001; 2001US-00927796.  
16-AUG-2001; 2001US-00931836.  
19-DEC-2001; 2001US-00028072.  
(GETH ) GENENTECH INC.  
Baker KP, Beresini M, Deforge L, Desnoyers L, Filvaroff E, Gao W;  
Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;  
Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;  
WPI; 2003-331925/31.  
N-PSDB; ACA04055.  
New secreted and transmembrane nucleic acids and polypeptides, designated  
as PRO, useful for treating inflammation, organ failure, atherosclerosis,  
cardiac injury, infertility, birth defects, premature aging, AIDS, or  
cancer.  
Claim 12; Fig 64; 659pp; English.  
The invention relates to an isolated nucleic acid comprising, or which is  
at least 80% identical to, or the full-length coding sequence of, any of  
the 275 nucleotide sequences, encoding the corresponding PRO polypeptide  
(one of 275 secreted or transmembrane proteins). The nucleic acid further  
comprises the full-length coding sequence of the DNA deposited under  
American Type Culture Collection (ATCC) accession number in a list given  
in the specification. Also included are vectors and host cells for  
producing PRO proteins, PRO fusion proteins, anti-PRO antibodies, PRO  
extracellular domains and mature sequences, methods of detecting PRO  
proteins, methods for stimulating the release of TNF-alpha (tumour  
necrosis factor alpha) from human blood, (and the proliferation of  
differentiation of chondrocyte cells, the proliferation of, or gene



CC	expression in pericyte cells, the release or proteoglycans from	KW	Tumour necrosis factor alpha release; TNF-alpha release;
CC	cartilage, proliferation of inner ear utricular supporting cells, the	KW	glucose uptake modulator; FFA uptake modulator;
CC	proliferation of T-lymphocyte cells, the release of a cytokine from	KW	cell proliferation stimulator; cell differentiation stimulator;
CC	peripheral blood mononuclear cells (PBMC), or the proliferation of	KW	cell differentiation inhibitor; cytokine release stimulator; tumour;
CC	endothelial cells), a method for modulating the uptake of glucose or free	KW	lung tumour; colon tumour; breast tumour; prostate tumour; rectal tumour;
CC	fatty acid (FFA) by skeletal muscle cells, a method for inhibiting the	KW	cervical tumour; liver tumour; chromosome mapping; gene mapping;
CC	binding of A-peptide to factor VIIA, or the differentiation of adipocyte	KW	gene therapy; chromosome identification; chromosome marker.
CC	cells, a method for detecting the presence of a tumour in a mammal and an	XX	
CC	oligonucleotide probe derived from any of the nucleotide sequences cited	OS	Homo sapiens.
CC	above. The nucleic acids and polypeptides are useful for treating	XX	
CC	inflammatory diseases, organ failure, premature aging, AIDS (acquired	PN	US2003022328-A1.
CC	infertility, birth defects, cancer, or diabetic complications. The	XX	30-JAN-2003.
CC	immunodeficiency syndrome), cancer, or diabetic complications. The	PD	
CC	nucleic acids are useful as hybridisation probes, in chromosome and gene	XX	
CC	mapping, and in generating antisense RNA or DNA. The polypeptides are	PF	16-APR-2002; 2002US-00123904.
CC	useful as pharmaceuticals, diagnostics, biosensors or bioreactors. Both	XX	31-MAR-1997; 97WO-US005230.
CC	are useful in tissue typing. The present sequence represents a PRO	PR	12-JUN-1998; 98WO-US012456.
XX	protein of the invention	PR	14-JUL-1998; 98WO-US014552.
SQ	Sequence 504 AA;	PR	28-AUG-1998; 98WO-US017888.
	Query Match 100.0%; Score 2615; DB 6; Length 504;	PR	10-SEP-1998; 98WO-US018824.
	Best Local Similarity 100.0%; Pred. No. 1.8e-258;	PR	14-SEP-1998; 98WO-US019093.
	Matches 504; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	PR	14-SEP-1998; 98WO-US019094.
QY	1 MEATGTWALLALALILLLTALSCTRARGHLPPOGTPPLILGNLLQLRPGALYSLMRL 60	PR	14-SEP-1998; 98WO-US019177.
DB	1 MEATGTWALLALALILLLTALSCTRARGHLPPOGTPPLILGNLLQLRPGALYSLMRL 60	PR	16-SEP-1998; 98WO-US019330.
QY	61 SKKYGPFVYIYIGMPRPVVVLVGQEAVERALGQAEFSGRGTVAMLEGTDFDGHGVFFSN 120	PR	17-SEP-1998; 98WO-US019437.
DB	61 SKKYGPFVYIYIGMPRPVVVLVGQEAVERALGQAEFSGRGTVAMLEGTDFDGHGVFFSN 120	PR	07-OCT-1998; 98WO-US021141.
QY	121 GERWRQLRFTMLALDLGKREGELIQAERCLIVETFOGTGPRFPDPSLLAQATSN 180	PR	29-OCT-1998; 98WO-US022991.
DB	121 GERWRQLRFTMLALDLGKREGELIQAERCLIVETFOGTGPRFPDPSLLAQATSN 180	PR	29-OCT-1998; 98WO-US022992.
QY	181 VVCSLLFGLRFSYEDKEFQAVVRAAGTLLGVSSQGGQTYEMPSWFLRPLPGPHKQLLHH 240	PR	20-NOV-1998; 98WO-US024855.
DB	181 VVCSLLFGLRFSYEDKEFQAVVRAAGTLLGVSSQGGQTYEMPSWFLRPLPGPHKQLLHH 240	PR	01-DEC-1998; 98WO-US025108.
QY	241 VSTLAFTVQVQHQHNLDAAGPARDLVDAFLKKAQEEQNPTETNKNMLMTVLYL 300	PR	05-JAN-1999; 98WO-US000106.
DB	241 VSTLAFTVQVQHQHNLDAAGPARDLVDAFLKKAQEEQNPTETNKNMLMTVLYL 300	PR	08-MAR-1999; 98WO-US005028.
QY	301 FAGTMTVSTTVGYTLALLMKYPHVQKWRBELNRELGAQAPSLGDRTRLPYTDVILHEA 360	PR	10-MAR-1999; 98WO-US005190.
DB	301 FAGTMTVSTTVGYTLALLMKYPHVQKWRBELNRELGAQAPSLGDRTRLPYTDVILHEA 360	PR	20-APR-1999; 98WO-US008615.
QY	361 QRLIALVPMGIPRTLMRTTRFRGYTLPGQTEVFPLLSILHDPNIFKHPEEFNPDRLDA 420	PR	14-MAY-1999; 98WO-US010733.
DB	361 QRLIALVPMGIPRTLMRTTRFRGYTLPGQTEVFPLLSILHDPNIFKHPEEFNPDRLDA 420	PR	02-JUN-1999; 98WO-US012552.
QY	421 DGRFRKHEAFPLSGKRVCLGEGAKAEFLFPTTILQAFSLSPCPDPTLSLKPTVSG 480	PR	01-SEP-1999; 98WO-US020111.
DB	421 DGRFRKHEAFPLSGKRVCLGEGAKAEFLFPTTILQAFSLSPCPDPTLSLKPTVSG 480	PR	08-SEP-1999; 98WO-US020594.
QY	481 LFNIPPAFQLOVRPDLHLSTTQTR 504	PR	13-SEP-1999; 98WO-US020944.
DB	481 LFNIPPAFQLOVRPDLHLSTTQTR 504	PR	15-SEP-1999; 98WO-US021090.
RESULT 11		PR	98WO-US021547.
ADA45583	ADA45583 standard; protein; 504 AA.	PR	05-OCT-1999; 98WO-US023089.
XX	AC	PR	29-NOV-1999; 98WO-US028214.
XX	AC	PR	30-NOV-1999; 98WO-US028313.
XX	DT	PR	30-NOV-1999; 98WO-US028409.
XX	DT	PR	01-DEC-1999; 98WO-US028301.
XX	DE	PR	01-DEC-1999; 98WO-US028634.
XX	DE	PR	02-DEC-1999; 98WO-US028551.
XX	DE	PR	02-DEC-1999; 98WO-US028565.
XX	DE	PR	16-DEC-1999; 98WO-US030095.
XX	DE	PR	20-DEC-1999; 98WO-US030911.
XX	DE	PR	20-DEC-1999; 98WO-US030999.
XX	DE	PR	30-DEC-1999; 98WO-US030720.
XX	DE	PR	30-DEC-1999; 98WO-US031243.
XX	DE	PR	30-DEC-1999; 98WO-US031274.
XX	DE	PR	2000WO-US000219.
XX	DE	PR	05-JAN-2000; 2000WO-US000277.
XX	DE	PR	06-JAN-2000; 2000WO-US000376.
XX	DE	PR	11-FEB-2000; 2000WO-US003565.
XX	DE	PR	18-FEB-2000; 2000WO-US004341.
XX	DE	PR	18-FEB-2000; 2000WO-US004342.
XX	DE	PR	22-FEB-2000; 2000WO-US004414.
XX	DE	PR	24-FEB-2000; 2000WO-US004914.
XX	DE	PR	24-FEB-2000; 2000WO-US005004.
XX	DE	PR	01-MAR-2000; 2000WO-US005601.
XX	DE	PR	02-MAR-2000; 2000WO-US005746.
XX	DE	PR	10-MAR-2000; 2000WO-US005841.
XX	DE	PR	15-MAR-2000; 2000WO-US006319.
XX	DE	PR	20-MAR-2000; 2000WO-US006884.
XX	DE	PR	20-MAR-2000; 2000WO-US007377.
XX	DE	PR	21-MAR-2000; 2000WO-US007532.

PR 30-MAR-2000; 2000WO-US008439.  
 PR 17-MAY-2000; 2000WO-US013705.  
 PR 22-MAY-2000; 2000WO-US014042.  
 PR 30-MAY-2000; 2000WO-US014941.  
 PR 02-JUN-2000; 2000WO-US015264.  
 PR 28-JUL-2000; 2000WO-US020710.  
 PR 11-AUG-2000; 2000WO-US022031.  
 PR 23-AUG-2000; 2000WO-US023522.  
 PR 24-AUG-2000; 2000WO-US023328.  
 PR 08-NOV-2000; 2000WO-US030952.  
 PR 01-DEC-2000; 2000WO-US030873.  
 PR 10-DEC-2000; 2000WO-US022678.  
 PR 20-DEC-2000; 2000US-00747259.  
 PR 20-DEC-2000; 2000WO-US034956.  
 PR 28-FEB-2001; 2001US-00796498.  
 PR 28-FEB-2001; 2001WO-US006520.  
 PR 01-MAR-2001; 2001WO-US006666.  
 PR 09-MAR-2001; 2001US-00802706.  
 PR 14-MAR-2001; 2001US-00808689.  
 PR 22-MAR-2001; 2001US-00816744.  
 PR 05-APR-2001; 2001US-00828366.  
 PR 10-MAY-2001; 2001US-00854208.  
 PR 10-MAY-2001; 2001US-00854280.  
 PR 18-MAY-2001; 2001US-00860216.  
 PR 25-MAY-2001; 2001US-00866034.  
 PR 25-MAY-2001; 2001WO-US017092.  
 PR 01-JUN-2001; 2001US-00872035.  
 PR 01-JUN-2001; 2001WO-US017800.  
 PR 05-JUN-2001; 2001US-00874503.  
 PR 14-JUN-2001; 2001US-00882636.  
 PR 19-JUN-2001; 2001US-00886342.  
 PR 20-JUN-2001; 2001WO-US019692.  
 PR 21-JUN-2001; 2001US-00887879.  
 PR 22-JUN-2001; 2001WO-US020116.  
 PR 29-JUN-2001; 2001WO-US021066.  
 PR 03-JUL-2001; 2001WO-US021735.  
 PR 18-JUL-2001; 2001US-00908827.  
 PR 06-AUG-2001; 2001US-00924419.  
 PR 09-AUG-2001; 2001US-00927796.  
 PR 16-AUG-2001; 2001US-00931836.  
 PR 19-DEC-2001; 2001US-00028072.

(GETH ) GENENTECH INC.

PI Baker KP, Beresini M, Deforge L, Desnoyers L, Filvaroff E, Gao W;  
 PI Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;  
 PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;

XX WPI; 2003-584997/55.  
 DR N-PSDB; ADA45582.

PT Novel secreted and transmembrane polypeptide for modulating biological  
 PT activity of cell expressing the polypeptide, identifying agonists or  
 PT antagonists of polypeptide, and as molecular weight markers.

XX Claim 12; Fig 64; 659pp; English.

CC The invention describes 305 nucleic acids encoding PRO (secreted and  
 CC transmembrane) polypeptides (I). (I) is useful for stimulating the  
 CC release of TNF-alpha from human blood, for modulating the uptake of  
 CC glucose or FFA by skeletal muscle cells or adipocyte cells, for  
 CC stimulating the proliferation or differentiation of chondrocyte cells,  
 CC cells, for stimulating the proliferation of or gene expression in pericyte  
 CC cells, for stimulating the release of proteoglycans from cartilage, for  
 CC stimulating the proliferation of inner ear utricular supporting cells,  
 CC the release of a cytokine from T-lymphocyte cells, for stimulating  
 CC A-peptide to factor VIIA, for inhibiting the differentiation of adipocyte  
 CC cells, for stimulating proliferation of endothelial cells, for detecting  
 CC the presence of tumour in a mammal. The tumour is lung, colon, breast,  
 CC prostate, rectal, cervical or liver tumour. The oligonucleotide probes  
 CC are useful for isolating genomic and cDNA nucleotide sequences or

CC antisense probes. (I) is also useful as therapeutic agent. PRO is useful  
 CC in assays to identify other proteins or molecules involved in binding  
 CC interaction. A polynucleotide (II) encoding (I) is useful in chromosome  
 CC and gene mapping, in generation of antisense RNA and DNA, in the  
 CC preparation of PRO polypeptide, for generating transgenic animals or  
 CC knockout animals which in turn are useful in the development and  
 CC screening of therapeutically useful reagents, in gene therapy, for  
 CC chromosome identification, as chromosome marker, and for generating  
 CC probes. An anti-(I)-antibody is useful in diagnostic assays for PRO. e.g.  
 CC detecting its expression in specific cells, tissues or serum, and for  
 CC affinity purification of PRO from recombinant cell culture or natural  
 CC sources. (I) and (II) are useful for tissue typing. This is the amino  
 CC acid sequence of a novel human secreted and transmembrane PRO  
 CC polypeptide.

XX Sequence 504 AA;

Query Match 100.0%; Score 2615; DB 6; Length 504;  
 Best Local Similarity 100.0%; Pred. No. 1.8e-258;  
 Matches 504; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEATGTWALLLALALLLLLTALSGTRARGHLPGPTPLPLGNLLQLRPGALYSGLMRL 60  
 Db 1 MEATGTWALLLALALLLLLTALSGTRARGHLPGPTPLPLGNLLQLRPGALYSGLMRL 60  
 QY 61 SKYGPVFTIYLGPMRPVVLVGQEAVERALGGQAEFFSGRTVAMLEGTDFDGHGVFNS 120  
 Db 61 SKYGPVFTIYLGPMRPVVLVGQEAVERALGGQAEFFSGRTVAMLEGTDFDGHGVFNS 120  
 QY 121 GERWQLRKFTMLALRDGLMGKREGEELIQAEARCLVETFGTEGRFPDPSLLAQATSN 180  
 Db 121 GERWQLRKFTMLALRDGLMGKREGEELIQAEARCLVETFGTEGRFPDPSLLAQATSN 180  
 QY 181 VVCSLLFGLRFSYEDKEFOAVRAAGGTLGVSSQGGQTYEMFSWFLRPLPGPHKQLLHH 240  
 Db 181 VVCSLLFGLRFSYEDKEFOAVRAAGGTLGVSSQGGQTYEMFSWFLRPLPGPHKQLLHH 240  
 QY 241 VSTLAFTVRQVQHQGNLDASGPARDLVDAFLLLKMAQESONPCTFTNKNMLMTVIYLL 300  
 Db 241 VSTLAFTVRQVQHQGNLDASGPARDLVDAFLLLKMAQESONPCTFTNKNMLMTVIYLL 300  
 QY 301 FAGTMVSTIVGYTILLMLKYPHVQKVRRELNRGLGAGQAPSLGDRTRLPYTDVILHEA 360  
 Db 301 FAGTMVSTIVGYTILLMLKYPHVQKVRRELNRGLGAGQAPSLGDRTRLPYTDVILHEA 360  
 QY 361 QRLLLVPMGIPRTLMRTTFRGTYTLPGQTEVEPLGSLIHDENIFKHPEEPNDRFLDA 420  
 Db 361 QRLLLVPMGIPRTLMRTTFRGTYTLPGQTEVEPLGSLIHDENIFKHPEEPNDRFLDA 420  
 QY 421 DGRFRKHEAFLPFLSKRVCLGEGAKAELELFFFTTILQAFSLSPCPDPTLSLKPTVSG 480  
 Db 421 DGRFRKHEAFLPFLSKRVCLGEGAKAELELFFFTTILQAFSLSPCPDPTLSLKPTVSG 480  
 QY 481 LFNIPPAFOLQVRPTDLHSTTQTR 504  
 Db 481 LFNIPPAFOLQVRPTDLHSTTQTR 504

RESULT 12

ADA76014  
 ID ADA76014 standard; protein; 504 AA.

XX ADA76014;  
 AC

XX 20-NOV-2003 (first entry)  
 DT

XX Human PRO polypeptide #32.  
 DE

XX Human; PRO; secreted polypeptide; transmembrane polypeptide;  
 KW tumour necrosis factor-alpha; TNF-alpha; chondrocyte cell; tumour;  
 KW cancer; adrenal; lung; colon; breast; prostate; rectum; kidney; cervix;  
 KW liver; microvascular endothelial cell; glucose; FFA;  
 KW skeletal muscle cell; adipocyte cell; pericyte cell;

Mon Sep 20 09:09:24 2004

us-10-669-693-2.rag

KW inner ear utricular supporting cell; T-lymphocyte cell;  
 KW endothelial cell tube formation; bone disorder; cartilage disorder;  
 KW sports injury; proteoglycan; articular cartilage defect; osteoarthritis;  
 KW rheumatoid arthritis; haemoglobin-associated disorder thalassaemia;  
 KW immune system cell infiltration.  
 XX  
 OS Homo sapiens.  
 XX  
 XX US2003073212-A1.  
 PN 17-APR-2003.  
 PD  
 XX  
 XX 16-APR-2002; 2002US-00123903.  
 PF  
 XX 31-MAR-1997; 97WO-US005230.  
 PR 12-JUN-1998; 98WO-US012456.  
 PR 14-JUL-1998; 98WO-US014552.  
 PR 28-AUG-1998; 98WO-US017888.  
 PR 10-SEP-1998; 98WO-US018824.  
 PR 14-SEP-1998; 98WO-US019093.  
 PR 14-SEP-1998; 98WO-US019094.  
 PR 14-SEP-1998; 98WO-US019177.  
 PR 16-SEP-1998; 98WO-US019330.  
 PR 17-SEP-1998; 98WO-US019437.  
 PR 07-OCT-1998; 98WO-US021141.  
 PR 29-OCT-1998; 98WO-US022991.  
 PR 29-OCT-1998; 98WO-US022992.  
 PR 29-OCT-1998; 98WO-US024855.  
 PR 01-DEC-1998; 98WO-US025108.  
 PR 05-JAN-1999; 99WO-US000108.  
 PR 08-MAR-1999; 99WO-US005190.  
 PR 10-MAR-1999; 99WO-US008615.  
 PR 14-MAY-1999; 99WO-US010733.  
 PR 02-JUN-1999; 99WO-US021252.  
 PR 01-SEP-1999; 99WO-US020111.  
 PR 08-SEP-1999; 99WO-US020594.  
 PR 13-SEP-1999; 99WO-US020944.  
 PR 15-SEP-1999; 99WO-US021090.  
 PR 15-SEP-1999; 99WO-US021547.  
 PR 05-OCT-1999; 99WO-US023089.  
 PR 29-NOV-1999; 99WO-US028214.  
 PR 30-NOV-1999; 99WO-US028313.  
 PR 30-NOV-1999; 99WO-US028409.  
 PR 01-DEC-1999; 99WO-US028301.  
 PR 01-DEC-1999; 99WO-US028634.  
 PR 02-DEC-1999; 99WO-US028551.  
 PR 02-DEC-1999; 99WO-US028564.  
 PR 02-DEC-1999; 99WO-US028565.  
 PR 16-DEC-1999; 99WO-US030095.  
 PR 20-DEC-1999; 99WO-US030911.  
 PR 20-DEC-1999; 99WO-US030999.  
 PR 22-DEC-1999; 99WO-US030720.  
 PR 30-DEC-1999; 99WO-US031243.  
 PR 30-DEC-1999; 99WO-US031274.  
 PR 05-JAN-2000; 2000WO-US000219.  
 PR 06-JAN-2000; 2000WO-US000277.  
 PR 06-JAN-2000; 2000WO-US000376.  
 PR 11-FEB-2000; 2000WO-US003565.  
 PR 18-FEB-2000; 2000WO-US004341.  
 PR 18-FEB-2000; 2000WO-US004342.  
 PR 22-FEB-2000; 2000WO-US004414.  
 PR 24-FEB-2000; 2000WO-US005004.  
 PR 24-FEB-2000; 2000WO-US005009.  
 PR 01-MAR-2000; 2000WO-US005601.  
 PR 02-MAR-2000; 2000WO-US005746.  
 PR 10-MAR-2000; 2000WO-US006319.  
 PR 15-MAR-2000; 2000WO-US006884.  
 PR 20-MAR-2000; 2000WO-US007377.  
 PR 21-MAR-2000; 2000WO-US007532.  
 PR 30-MAR-2000; 2000WO-US008439.  
 PR 17-MAY-2000; 2000WO-US013705.  
 PR 22-MAY-2000; 2000WO-US014042.  
 PR 30-MAY-2000; 2000WO-US014941.  
 PR 02-JUN-2000; 2000WO-US015264.  
 PR 28-JUN-2000; 2000WO-US020710.  
 PR 11-AUG-2000; 2000WO-US022031.  
 PR 23-AUG-2000; 2000WO-US023522.  
 PR 24-AUG-2000; 2000WO-US023328.  
 PR 08-NOV-2000; 2000WO-US030952.  
 PR 10-NOV-2000; 2000WO-US030873.  
 PR 01-DEC-2000; 2000WO-US032678.  
 PR 20-DEC-2000; 2000US-00747259.  
 PR 20-DEC-2000; 2000WO-US034956.  
 PR 28-FEB-2001; 2001US-00796498.  
 PR 28-FEB-2001; 2001WO-US006520.  
 PR 01-MAR-2001; 2001WO-US006666.  
 PR 09-MAR-2001; 2001US-00802706.  
 PR 14-MAR-2001; 2001US-00808689.  
 PR 22-MAR-2001; 2001US-00816744.  
 PR 05-APR-2001; 2001US-00828366.  
 PR 10-MAY-2001; 2001US-00854208.  
 PR 10-MAY-2001; 2001US-00854280.  
 PR 18-MAY-2001; 2001US-00860216.  
 PR 25-MAY-2001; 2001US-00866028.  
 PR 25-MAY-2001; 2001US-00866034.  
 PR 25-MAY-2001; 2001WO-US017092.  
 PR 01-JUN-2001; 2001US-00872035.  
 PR 01-JUN-2001; 2001WO-US017800.  
 PR 05-JUN-2001; 2001US-00874503.  
 PR 14-JUN-2001; 2001US-00882636.  
 PR 19-JUN-2001; 2001US-00886342.  
 PR 20-JUN-2001; 2001WO-US019692.  
 PR 21-JUN-2001; 2001US-00887879.  
 PR 22-JUN-2001; 2001WO-US020116.  
 PR 29-JUN-2001; 2001WO-US021066.  
 PR 09-JUL-2001; 2001WO-US021735.  
 PR 18-JUL-2001; 2001US-00908827.  
 PR 06-AUG-2001; 2001US-00924419.  
 PR 09-AUG-2001; 2001US-00927796.  
 PR 16-AUG-2001; 2001US-00931836.  
 PR 19-DEC-2001; 2001US-00028072.  
 XX  
 PA (GETH ) GENENTECH INC.  
 XX Baker KP, Beresini M, Deforge L, Desnoyers L, Filvaroff E, Gao W;  
 PI Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;  
 PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;  
 XX WPI; 2003-687639/65.  
 DR N-PSDB; ADA76013.  
 DR  
 XX New isolated nucleic acid encoding a secreted and transmembrane  
 PT polypeptide, designated e.g. PRO1114 or PRO4978, useful in chromosome and  
 PT gene mapping, in generating antisense RNA and DNA, and in gene therapy.  
 XX  
 PS Claim 12; Fig 64; 659pp; English.  
 XX  
 CC The invention relates to isolated human PRO polypeptides (secreted and  
 CC transmembrane polypeptides) and the polynucleotides encoding them. The  
 CC invention also relates to an antibody which specifically binds to a PRO  
 CC polypeptide, a method for stimulating the release of tumour necrosis  
 CC factor-alpha (TNF-alpha) from human blood, a method for stimulating the  
 CC proliferation or differentiation of chondrocyte cells and a method for  
 CC detecting the presence of a tumour in a mammal (e.g. adrenal, lung,  
 CC colon, breast, prostate, rectal, kidney, cervical and liver tumours). The  
 CC polynucleotides are useful in molecular biology, including uses as  
 CC hybridisation probes, in chromosome and gene mapping, in generating  
 CC antisense RNA and DNA and in gene therapy. The polynucleotides may also  
 CC be used in preparing PRO polypeptides by recombinant techniques and in  
 CC generating either transgenic animals or knock-out animals which are  
 CC useful in the development and screening of therapeutically useful  
 CC reagents. The PRO polypeptides or antibodies are used in preparing a  
 CC medicament for treating a condition responsive to the polypeptides or  
 CC antibodies, such as tumours, for stimulating and inhibiting proliferation

CC of human microvascular endothelial cells, for modulating the uptake of  
CC glucose or FFA by skeletal muscle cells or adipocyte cells, for  
CC stimulating differentiation of adipocyte cells, for stimulating  
CC proliferation of or gene expression in pericyte cells, for stimulating  
CC the proliferation of inner ear utricular supporting cells or T-lymphocyte  
CC cells, for inducing endothelial cell tube formation and for treating  
CC various bone and/or cartilage disorders such as sports injuries and  
CC arthritis. PRO polypeptides which stimulate the release of proteoglycans  
CC from cartilage are useful for treating sports-related joint problems, PRO  
CC articular cartilage defects, osteoarthritis and rheumatoid arthritis. PRO  
CC polypeptides are also useful for treating various mammalian haemoglobin-  
CC associated disorders such as various thalassaemias and conditions which  
CC may benefit from enhanced local immune system cell infiltration. This  
CC sequence represents a human PRO polypeptide of the invention. Note: The  
CC sequence data for this patent is also available in electronic format from  
CC USPTO at [seqdata.uspto.gov/sequence.html](http://seqdata.uspto.gov/sequence.html).  
XX  
SQ Sequence 504 AA;

Query Match 100.0%; Score 2615; DB 6; Length 504;  
Best Local Similarity 100.0%; Pred. No. 1.8e-258;  
Matches 504; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEATGTWALLLALLLTLALSGTRARGHLPGPTPLPLGNNLQLRPGALYSGLMRL 60  
DB 1 MEATGTWALLLALLLTLALSGTRARGHLPGPTPLPLGNNLQLRPGALYSGLMRL 60

QY 61 SKKYGPFVTVLGPWRPVVLVGOAVREALGQAEFSGRGTVAMLEGTFDGHGVFFSN 120  
DB 61 SKKYGPFVTVLGPWRPVVLVGOAVREALGQAEFSGRGTVAMLEGTFDGHGVFFSN 120

QY 121 GERWRQLRKFTMLALRDLGMKRGEEELIOAEARCLVETFGTEGRPDPSLLLAQATSN 180  
DB 121 GERWRQLRKFTMLALRDLGMKRGEEELIOAEARCLVETFGTEGRPDPSLLLAQATSN 180

QY 181 VVCSLLPLRFSTEDKEFQAVVRAAGTLLGVSSQGGQTYEMFSWFLRPLPGPHKQLLHH 240  
DB 181 VVCSLLPLRFSTEDKEFQAVVRAAGTLLGVSSQGGQTYEMFSWFLRPLPGPHKQLLHH 240

QY 241 VSTLAFTVQVQHQNLDSAGPARDLVDAFLIKMAQEQNPCTETNNMLMTVIYLL 300  
DB 241 VSTLAFTVQVQHQNLDSAGPARDLVDAFLIKMAQEQNPCTETNNMLMTVIYLL 300

QY 301 FAGTMVSTTVGTYLLLMKYPHQVQWVREELNRELGAQAPSLGDRTRLPYTDVILHEA 360  
DB 301 FAGTMVSTTVGTYLLLMKYPHQVQWVREELNRELGAQAPSLGDRTRLPYTDVILHEA 360

QY 361 ORLLALVPMGIPRLMRTTRFRGYTLPQGTVEVPLLSILHDPNIFKHPEEFNDRFLDA 420  
DB 361 ORLLALVPMGIPRLMRTTRFRGYTLPQGTVEVPLLSILHDPNIFKHPEEFNDRFLDA 420

QY 421 DGRFRKHEAFPLPSLCKRVCLGSLAKAELEFPTTILQAFSLSECPDPTLSLKTQVSG 480  
DB 421 DGRFRKHEAFPLPSLCKRVCLGSLAKAELEFPTTILQAFSLSECPDPTLSLKTQVSG 480

QY 481 LENIPPAFQLQVRPTDLHSTTQTR 504  
DB 481 LENIPPAFQLQVRPTDLHSTTQTR 504

## RESULT 13

ADA18664  
ID ADA18664 standard; protein; 504 AA.  
XX  
AC ADA18664;  
XX  
DT 20-NOV-2003 (first entry)  
XX  
DE Human PRO polypeptide #32.  
XX  
KW Human; PRO: secreted polypeptide; transmembrane polypeptide;  
KW tumour necrosis factor-alpha; TNF-alpha; blood; chondrocyte cell; lung;  
KW colon; breast; prostate; rectum; cervix; liver; tumour; cancer;

KW glucose uptake; FFA; adipocyte cell; pericyte cell; proteoglycan;  
KW cartilage; inner ear utricular supporting cell; cytokine; A-peptide;  
KW factor VIIA; endothelial cell.  
OS Homo sapiens.  
XX US2003054517-A1.  
XX  
PD 20-MAR-2003.  
XX  
PF 08-MAY-2002; 2002US-00141755.  
XX  
PR 31-MAR-1997; 97WO-US005230.  
PR 12-JUN-1998; 98WO-US012456.  
PR 14-JUL-1998; 98WO-US014552.  
PR 28-AUG-1998; 98WO-US017888.  
PR 10-SEP-1998; 98WO-US018824.  
PR 14-SEP-1998; 98WO-US019093.  
PR 14-SEP-1998; 98WO-US019094.  
PR 14-SEP-1998; 98WO-US019177.  
PR 16-SEP-1998; 98WO-US019330.  
PR 17-SEP-1998; 98WO-US019437.  
PR 07-OCT-1998; 98WO-US021141.  
PR 29-OCT-1998; 98WO-US022591.  
PR 29-OCT-1998; 98WO-US022592.  
PR 20-NOV-1998; 98WO-US024855.  
PR 01-DEC-1998; 98WO-US025108.  
PR 05-JAN-1999; 98WO-US000106.  
PR 08-MAR-1999; 98WO-US005028.  
PR 10-MAR-1999; 98WO-US005190.  
PR 20-APR-1999; 98WO-US008615.  
PR 14-MAY-1999; 98WO-US010733.  
PR 02-JUN-1999; 98WO-US012252.  
PR 01-SEP-1999; 98WO-US020111.  
PR 08-SEP-1999; 98WO-US020594.  
PR 13-SEP-1999; 98WO-US020944.  
PR 15-SEP-1999; 98WO-US021090.  
PR 15-SEP-1999; 98WO-US021547.  
PR 05-OCT-1999; 98WO-US023089.  
PR 29-NOV-1999; 98WO-US028214.  
PR 30-NOV-1999; 98WO-US028313.  
PR 30-NOV-1999; 98WO-US028409.  
PR 01-DEC-1999; 98WO-US028301.  
PR 01-DEC-1999; 98WO-US028634.  
PR 02-DEC-1999; 98WO-US028551.  
PR 02-DEC-1999; 98WO-US028564.  
PR 02-DEC-1999; 98WO-US028565.  
PR 16-DEC-1999; 98WO-US030095.  
PR 20-DEC-1999; 98WO-US030911.  
PR 20-DEC-1999; 98WO-US030999.  
PR 22-DEC-1999; 98WO-US030720.  
PR 30-DEC-1999; 98WO-US031243.  
PR 30-DEC-1999; 98WO-US031274.  
PR 05-JAN-2000; 2000WO-US000219.  
PR 06-JAN-2000; 2000WO-US000277.  
PR 06-JAN-2000; 2000WO-US000376.  
PR 11-FEB-2000; 2000WO-US003565.  
PR 18-FEB-2000; 2000WO-US004341.  
PR 18-FEB-2000; 2000WO-US004342.  
PR 22-FEB-2000; 2000WO-US004414.  
PR 24-FEB-2000; 2000WO-US004914.  
PR 24-FEB-2000; 2000WO-US005004.  
PR 01-MAR-2000; 2000WO-US005601.  
PR 02-MAR-2000; 2000WO-US005746.  
PR 02-MAR-2000; 2000WO-US005841.  
PR 10-MAR-2000; 2000WO-US006319.  
PR 15-MAR-2000; 2000WO-US006884.  
PR 20-MAR-2000; 2000WO-US007377.  
PR 21-MAR-2000; 2000WO-US007532.  
PR 30-MAR-2000; 2000WO-US008439.  
PR 17-MAY-2000; 2000WO-US013705.  
PR 22-MAY-2000; 2000WO-US014042.  
PR 30-MAY-2000; 2000WO-US014941.

CC supporting cells, for stimulating the release of cytokines from PBMC  
CC cells, for inhibiting the binding of A-peptide to factor VIIA, for  
CC inhibiting the differentiation of adipocyte cells and for stimulating the  
CC proliferation of endothelial cells. This sequence represents a human PRO  
CC polypeptide of the invention. Note: The sequence data for this patent is  
CC also available in electronic format from USPTO at  
CC seqdata.uspto.gov/sequence.html.  
XX  
SQ Sequence 504 AA;  
  
Query Match 100.0%; Score 2615; DB 6; Length 504;  
Best Local Similarity 100.0%; Pred. No. 1.8e-258; Indels 0; Gaps 0;  
Matches 504; Conservative 0; Mismatches 0;  
  
QY 1 MEATGTWALLALALLLLTLLSALSGTRARGHLPGQPTPLLLGNLQLRPGALYGLML 60  
DB 1 MEATGTWALLALALLLLTLLSALSGTRARGHLPGQPTPLLLGNLQLRPGALYGLML 60  
  
QY 61 SKKYGPFVTIYLGPRPVVVLVGQAVREALGGQAEFSGRTVAMLEGTDFGHGVFFSN 120  
DB 61 SKKYGPFVTIYLGPRPVVVLVGQAVREALGGQAEFSGRTVAMLEGTDFGHGVFFSN 120  
  
QY 121 GERWQLRKFTMLALRDLMGKREGEELIQAEARCLVETFOCTEGRPDPFSLLLAQATSN 180  
DB 121 GERWQLRKFTMLALRDLMGKREGEELIQAEARCLVETFOCTEGRPDPFSLLLAQATSN 180  
  
QY 181 VVCSLLFGLRFSYEDKEFOAVVRAAGTLLGVSSGGQTYEMFSWFLRPLPGPHKQLLHH 240  
DB 181 VVCSLLFGLRFSYEDKEFOAVVRAAGTLLGVSSGGQTYEMFSWFLRPLPGPHKQLLHH 240  
  
QY 241 VSTLAAFTVRQVQHQGNLDASGPARDLVDAFLKMAQEEQNGTEFTNKMLMTVIYLL 300  
DB 241 VSTLAAFTVRQVQHQGNLDASGPARDLVDAFLKMAQEEQNGTEFTNKMLMTVIYLL 300  
  
QY 301 FAGTMTVSTVGYTLLMLMKYDHVKWVREELNRELGAGQAPSLGDRTRLPTVDVHLHEA 360  
DB 301 FAGTMTVSTVGYTLLMLMKYDHVKWVREELNRELGAGQAPSLGDRTRLPTVDVHLHEA 360  
  
QY 361 QRLALLVPMGIPRTLMRTTRFRGYTLPOGTEVFPPLGSLHDPNIFKHPFEENPDRFLDA 420  
DB 361 QRLALLVPMGIPRTLMRTTRFRGYTLPOGTEVFPPLGSLHDPNIFKHPFEENPDRFLDA 420  
  
QY 421 DGRFRKHEAPLPSLGKRVCLGEGKAKELFLPFTTILQAFSLSPDPTLSLKPTVSG 480  
DB 421 DGRFRKHEAPLPSLGKRVCLGEGKAKELFLPFTTILQAFSLSPDPTLSLKPTVSG 480  
  
QY 481 LFNIPPAFOLOVRPTDLHSTTQTR 504  
DB 481 LFNIPPAFOLOVRPTDLHSTTQTR 504  
  
RESULT 14  
ADA61287  
ID ADA61287 standard; protein; 504 AA.  
XX  
AC ADA61287;  
XX  
DT 20-NOV-2003 (first entry)  
XX  
DE Homo sapiens.  
XX  
KW Human; secreted and transmembrane protein; PRO;  
KW Tumour necrosis factor alpha release; TNF-alpha release;  
KW glucose uptake modulator; FFA uptake modulator;  
KW cell proliferation stimulator; cell differentiation stimulator;  
KW cell differentiation inhibitor; cytokine release stimulator; tumour;  
KW lung tumour; colon tumour; breast tumour; prostate tumour; rectal tumour;  
KW cervical tumour; liver tumour; chromosome mapping; gene mapping;  
KW gene therapy; chromosome identification; chromosome marker.  
XX  
OS Novel.  
OS human.  
OS secreted.

PR 02-JUN-2000; 2000WO-US015264.  
PR 28-JUL-2000; 2000WO-US020710.  
PR 11-AUG-2000; 2000WO-US022031.  
PR 23-AUG-2000; 2000WO-US023522.  
PR 24-AUG-2000; 2000WO-US023328.  
PR 08-NOV-2000; 2000WO-US030952.  
PR 10-NOV-2000; 2000WO-US030873.  
PR 01-DEC-2000; 2000WO-US032678.  
PR 20-DEC-2000; 2000US-00747259.  
PR 20-DEC-2000; 2000WO-US034956.  
PR 28-FEB-2001; 2001US-00796498.  
PR 28-FEB-2001; 2001WO-US006520.  
PR 01-MAR-2001; 2001WO-US006666.  
PR 09-MAR-2001; 2001US-00802706.  
PR 14-MAR-2001; 2001US-00808689.  
PR 22-MAR-2001; 2001US-00816744.  
PR 05-APR-2001; 2001US-00828366.  
PR 10-MAY-2001; 2001US-00854208.  
PR 10-MAY-2001; 2001US-00854280.  
PR 18-MAY-2001; 2001US-00860216.  
PR 25-MAY-2001; 2001US-00866028.  
PR 25-MAY-2001; 2001US-00866034.  
PR 25-MAY-2001; 2001WO-US017092.  
PR 01-JUN-2001; 2001US-00872035.  
PR 01-JUN-2001; 2001WO-US017800.  
PR 05-JUN-2001; 2001US-00874503.  
PR 14-JUN-2001; 2001US-00882636.  
PR 19-JUN-2001; 2001US-00886342.  
PR 20-JUN-2001; 2001WO-US019692.  
PR 21-JUN-2001; 2001US-00887879.  
PR 22-JUN-2001; 2001WO-US020116.  
PR 29-JUN-2001; 2001US-0021066.  
PR 09-JUL-2001; 2001WO-US021735.  
PR 18-JUL-2001; 2001US-00908827.  
PR 06-AUG-2001; 2001US-00924419.  
PR 09-AUG-2001; 2001US-00927796.  
PR 16-AUG-2001; 2001US-00931836.  
PR 19-DEC-2001; 2001US-00028072.  
XX  
PA (GETH ) GENENTECH INC.  
XX  
PI Baker KP, Beresini M, Deforge L, Deanoyers L, Filvaroff E, Gao W;  
PI Gerritsen WB, Goddard A, Godowski PJ, Gurney AL, Sherwood S;  
PI Smith V, Stewart TA, Tumas D, Wacanabe CK, Wood WI, Zhang Z;  
XX  
DR WPI; 2003-521854/49.  
DR N-PSDB; ADA18663.  
XX  
PT New PRO nucleic acid, useful for preparing a composition for treating  
PT e.g., tumors.  
XX  
PS Claim 12; Fig 64; 660pp; English.  
XX  
CC The invention relates to isolated human PRO polypeptides (secreted and  
CC transmembrane polypeptides) and the polynucleotides encoding them. The  
CC invention also relates to an antibody which specifically binds to a PRO  
CC polypeptide, a method for stimulating the release of tumour necrosis  
CC factor-alpha (TNF-alpha) from human blood, a method for stimulating the  
CC proliferation or differentiation of chondrocyte cells and a method for  
CC detecting the presence of a tumour in a mammal (e.g. lung, colon, breast,  
CC prostate, rectal, cervical and liver tumours). The polynucleotides are  
CC useful in molecular biology, including uses as hybridisation probes, in  
CC chromosome and gene mapping, in generating antisense RNA and DNA and in  
CC gene therapy. The polynucleotides may also be used in preparing PRO  
CC polypeptides by recombinant techniques and in generating either  
CC transgenic animals or knock-out animals which are useful in the  
CC development and screening of therapeutically useful reagents. The PRO  
CC polypeptides or antibodies are used in preparing a medicament for  
CC treating a condition responsive to the polypeptides or antibodies, such  
CC as tumours, for modulating the uptake of glucose or FFA by adipocyte  
CC cells, for stimulating the proliferation of or gene expression in  
CC pericyte cells, for stimulating the release of proteoglycans from  
CC cartilage, for stimulating the proliferation of inner ear utricular

OS and.  
OS transmembrane.  
OS protein.  
XX PRO1906.

PN US2003049816-A1.

XX

PD 13-MAR-2003.

XX 15-APR-2002; 2002US-00123262.

PR 31-MAR-1997; 97WO-US005230.

PR 12-JUN-1998; 98WO-US012456.

PR 28-AUG-1998; 98WO-US014552.

PR 10-SEP-1998; 98WO-US017888.

PR 14-SEP-1998; 98WO-US018824.

PR 14-SEP-1998; 98WO-US019093.

PR 14-SEP-1998; 98WO-US019094.

PR 16-SEP-1998; 98WO-US019177.

PR 17-SEP-1998; 98WO-US019330.

PR 07-OCT-1998; 98WO-US019437.

PR 29-OCT-1998; 98WO-US021141.

PR 29-OCT-1998; 98WO-US022991.

PR 20-NOV-1998; 98WO-US022992.

PR 01-DEC-1998; 98WO-US024855.

PR 05-JAN-1999; 98WO-US025108.

PR 08-MAR-1999; 98WO-US005028.

PR 10-MAR-1999; 98WO-US005190.

PR 20-APR-1999; 98WO-US008615.

PR 14-MAY-1999; 98WO-US010733.

PR 02-JUN-1999; 98WO-US012252.

PR 08-SEP-1999; 98WO-US020111.

PR 13-SEP-1999; 98WO-US020594.

PR 15-SEP-1999; 98WO-US020944.

PR 15-SEP-1999; 98WO-US021090.

PR 03-OCT-1999; 98WO-US021547.

PR 29-NOV-1999; 98WO-US023089.

PR 30-NOV-1999; 98WO-US028214.

PR 30-NOV-1999; 98WO-US028313.

PR 01-DEC-1999; 98WO-US028409.

PR 01-DEC-1999; 98WO-US028301.

PR 02-DEC-1999; 98WO-US028634.

PR 02-DEC-1999; 98WO-US028551.

PR 02-DEC-1999; 98WO-US028564.

PR 16-DEC-1999; 98WO-US028565.

PR 20-DEC-1999; 98WO-US030095.

PR 20-DEC-1999; 98WO-US030911.

PR 22-DEC-1999; 98WO-US030999.

PR 30-DEC-1999; 98WO-US030720.

PR 30-DEC-1999; 98WO-US031243.

PR 05-JAN-2000; 98WO-US031274.

PR 06-JAN-2000; 2000WO-US000219.

PR 06-JAN-2000; 2000WO-US000277.

PR 11-FEB-2000; 2000WO-US000376.

PR 18-FEB-2000; 2000WO-US003565.

PR 18-FEB-2000; 2000WO-US004341.

PR 22-FEB-2000; 2000WO-US004342.

PR 24-FEB-2000; 2000WO-US004414.

PR 24-FEB-2000; 2000WO-US004914.

PR 01-MAR-2000; 2000WO-US005004.

PR 02-MAR-2000; 2000WO-US005601.

PR 02-MAR-2000; 2000WO-US005746.

PR 10-NAR-2000; 2000WO-US006319.

PR 15-MAR-2000; 2000WO-US006884.

PR 20-MAR-2000; 2000WO-US007377.

PR 21-MAR-2000; 2000WO-US007532.

PR 30-MAR-2000; 2000WO-US008439.

PR 17-MAY-2000; 2000WO-US013705.

PR 22-MAY-2000; 2000WO-US014042.

PR 30-MAY-2000; 2000WO-US014941.

PR 02-JUN-2000; 2000WO-US015264.

PR 28-JUL-2000; 2000WO-US020710.

PR 11-AUG-2000; 2000WO-US022031.

PR 23-AUG-2000; 2000WO-US023522.

PR 24-AUG-2000; 2000WO-US023328.

PR 08-NOV-2000; 2000WO-US030952.

PR 10-NOV-2000; 2000WO-US030873.

PR 01-DEC-2000; 2000WO-US032678.

PR 20-DEC-2000; 2000US-00747259.

PR 20-DEC-2000; 2000WO-US034956.

PR 28-FEB-2001; 2001US-00796498.

PR 28-FEB-2001; 2001WO-US006520.

PR 01-MAR-2001; 2001WO-US006666.

PR 09-MAR-2001; 2001US-00802706.

PR 14-MAR-2001; 2001US-00808689.

PR 22-MAR-2001; 2001US-00816744.

PR 05-APR-2001; 2001US-00828366.

PR 10-MAY-2001; 2001US-00854208.

PR 10-MAY-2001; 2001US-00854280.

PR 18-MAY-2001; 2001US-00860216.

PR 25-MAY-2001; 2001US-00866028.

PR 25-MAY-2001; 2001US-00866034.

PR 25-MAY-2001; 2001WO-US017092.

PR 01-JUN-2001; 2001US-00872035.

PR 01-JUN-2001; 2001WO-US017800.

PR 05-JUN-2001; 2001US-00874503.

PR 14-JUN-2001; 2001US-00882636.

PR 19-JUN-2001; 2001US-00886342.

PR 20-JUN-2001; 2001WO-US019692.

PR 21-JUN-2001; 2001US-00887879.

PR 22-JUN-2001; 2001WO-US020116.

PR 29-JUN-2001; 2001WO-US021066.

PR 09-JUL-2001; 2001WO-US021735.

PR 18-JUL-2001; 2001US-00908827.

PR 06-AUG-2001; 2001US-00924419.

PR 09-AUG-2001; 2001US-00927796.

PR 16-AUG-2001; 2001US-00931836.

PR 19-DEC-2001; 2001US-00028072.

(GETH ) GENENTECH INC.

Baker KP, Heresini M, Deforge L, Desnoyers L, Filvaroff E, Gao W; Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S; Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z; WPI; 2003-695892/66.

N-PSDB; ADA61286.

New PRO nucleic acid and encode polypeptides, are useful for manufacturing a medicament for diagnosing or treating cancer.

Claim 12; Fig 64; 660pp; English.

The invention describes 305 nucleic acids encoding PRO (secreted and transmembrane) polypeptides (I). (I) is useful for stimulating the release of TNF-alpha from human blood, for modulating the uptake of glucose or FFA by skeletal muscle cells or adipocyte cells, for stimulating the proliferation or differentiation of chondrocyte cells, for stimulating the proliferation of or gene expression in pericyte cells, for stimulating the release of proteoglycans from cartilage, for stimulating the proliferation of inner ear utricular supporting cells, for stimulating the proliferation of T-lymphocyte cells, for stimulating the release of a cytokine from PMBC cells, for inhibiting the binding of A-peptide to factor VIIA, for inhibiting the differentiation of adipocyte cells, for stimulating proliferation of endothelial cells, for detecting the presence of tumour in a mammal. The tumour is lung, colon, breast, prostate, rectal, cervical or liver tumour. The oligonucleotide probes are useful for isolating genomic and cDNA nucleotide sequences or antisense probes. (I) is also useful as therapeutic agent. PRO is useful in assays to identify other proteins or molecules involved in binding interaction. A polynucleotide (II) encoding (I) is useful in chromosome and gene mapping in generation of antisense RNA and DNA, in the preparation of PRO polypeptide, for generating transgenic animals or knockout animals which in turn are useful in the development and

Mon Sep 20 09:09:24 2004

us-10-669-693-2.rag

CC screening of therapeutically useful reagents, in gene therapy, for  
CC chromosome identification, as chromosome marker, and for generating  
CC probes. An anti-(I)-antibody is useful in diagnostic assays for PRO, e.g.  
CC detecting its expression in specific cells, tissues or serum, and for  
CC affinity purification of PRO from recombinant cell culture or natural  
CC sources. (I) and (II) are useful for tissue typing. This is the amino  
CC acid sequence of a novel human secreted and transmembrane PRO  
CC polypeptide.

XX Sequence 504 AA;

Query Match 100.0%; Score 2615; DB 6; Length 504;  
Best Local Similarity 100.0%; Pred. No. 1.8e-258; Indels 0; Gaps 0;  
Matches 504; Conservative 0; Mismatches 0;

QY 1 MEATGTWALLALALLLLTALSTRARGHLPGPTPLPLGNLLQLRPGALYSGLMRL 60  
Db 1 MEATGTWALLALALLLLTALSTRARGHLPGPTPLPLGNLLQLRPGALYSGLMRL 60  
QY 61 SKKYGPFVTIYLGPRVVLVQBAVREALGGQAEBSGRTVAMLEGTDPGHGVFSN 120  
Db 61 SKKYGPFVTIYLGPRVVLVQBAVREALGGQAEBSGRTVAMLEGTDPGHGVFSN 120  
QY 121 GERWQLAKFTMLARDLGMKRGEBELIOAEARCLVETFOCTEGRPPDPSPSLLLAQATSN 180  
Db 121 GERWQLAKFTMLARDLGMKRGEBELIOAEARCLVETFOCTEGRPPDPSPSLLLAQATSN 180  
QY 181 VVCSLLFGLRFSYEDKEFAQVRAAGTLLGVSSQGGTYEMFSWFLRPLPGPHKQLLHH 240  
Db 181 VVCSLLFGLRFSYEDKEFAQVRAAGTLLGVSSQGGTYEMFSWFLRPLPGPHKQLLHH 240  
QY 241 VSTLAAFTVRQVQOQGNLDASGPARDLVDAFLKWAQBEQNGTEFTKNMLMTVIYLL 300  
Db 241 VSTLAAFTVRQVQOQGNLDASGPARDLVDAFLKWAQBEQNGTEFTKNMLMTVIYLL 300  
QY 301 FAGTWTVTTCYGTLLLMKYDHVKWREELNRELGAGQPSLGRTRPLPYTDVAVLHEA 360  
Db 301 FAGTWTVTTCYGTLLLMKYDHVKWREELNRELGAGQPSLGRTRPLPYTDVAVLHEA 360  
QY 361 QRLALLVPMGIPRTLMRTFRGYTLPGQTEVPPLLSILHDPNIFKHPHEEFPNDRFLDA 420  
Db 361 QRLALLVPMGIPRTLMRTFRGYTLPGQTEVPPLLSILHDPNIFKHPHEEFPNDRFLDA 420  
QY 421 DGRFRKHEAFLPFLGKRVCLGELAKAEFLFFTTILQAFSLSPCPDPTLSLKPTVSG 480  
Db 421 DGRFRKHEAFLPFLGKRVCLGELAKAEFLFFTTILQAFSLSPCPDPTLSLKPTVSG 480  
QY 481 LFNIPPAFOLOVRPTDLHSTTQIR 504  
Db 481 LFNIPPAFOLOVRPTDLHSTTQIR 504

RESULT 15

ADBL9072  
ID ADBL9072 standard; protein; 504 AA.

XX ADB19072;

XX 20-NOV-2003 (first entry)

XX Novel human secreted and transmembrane protein PRO1906.

XX Human; secreted and transmembrane protein; PRO;  
KW Tumour necrosis factor alpha release; TNF-alpha release;  
KW glucose uptake modulator; FFA uptake modulator;  
KW cell proliferation stimulator; cell differentiation stimulator;  
KW cell differentiation inhibitor; cytokine release.

XX Homo sapiens.

XX US2003068796-A1.

XX 10-APR-2003.

PD

XX 15-APR-2002; 2002US-00123261.  
PF 97WO-US005230.  
XX 31-MAR-1997; 98WO-US012456.  
PR 12-JUN-1998; 98WO-US014552.  
PR 14-JUL-1998; 98WO-US017888.  
PR 28-AUG-1998; 98WO-US018824.  
PR 10-SEP-1998; 98WO-US019093.  
PR 14-SEP-1998; 98WO-US019094.  
PR 14-SEP-1998; 98WO-US019177.  
PR 14-SEP-1998; 98WO-US019330.  
PR 16-SEP-1998; 98WO-US019437.  
PR 07-OCT-1998; 98WO-US021141.  
PR 29-OCT-1998; 98WO-US022991.  
PR 29-OCT-1998; 98WO-US022992.  
PR 20-NOV-1998; 98WO-US024855.  
PR 01-DEC-1998; 98WO-US025108.  
PR 05-JAN-1999; 99WO-US000106.  
PR 08-MAR-1999; 99WO-US005028.  
PR 10-MAR-1999; 99WO-US005190.  
PR 20-APR-1999; 99WO-US008615.  
PR 14-MAY-1999; 99WO-US010733.  
PR 02-JUN-1999; 99WO-US012252.  
PR 01-SEP-1999; 99WO-US020111.  
PR 08-SEP-1999; 99WO-US020594.  
PR 13-SEP-1999; 99WO-US020944.  
PR 15-SEP-1999; 99WO-US021090.  
PR 15-SEP-1999; 99WO-US021547.  
PR 05-OCT-1999; 99WO-US023089.  
PR 30-NOV-1999; 99WO-US028214.  
PR 30-NOV-1999; 99WO-US028313.  
PR 30-NOV-1999; 99WO-US028409.  
PR 01-DEC-1999; 99WO-US028301.  
PR 01-DEC-1999; 99WO-US028634.  
PR 02-DEC-1999; 99WO-US028551.  
PR 02-DEC-1999; 99WO-US028564.  
PR 16-DEC-1999; 99WO-US030095.  
PR 20-DEC-1999; 99WO-US030911.  
PR 20-DEC-1999; 99WO-US030999.  
PR 22-DEC-1999; 99WO-US030720.  
PR 30-DEC-1999; 99WO-US031243.  
PR 30-DEC-1999; 99WO-US031274.  
PR 06-JAN-2000; 2000WO-US000219.  
PR 05-JAN-2000; 2000WO-US000277.  
PR 06-JAN-2000; 2000WO-US000376.  
PR 11-FEB-2000; 2000WO-US003565.  
PR 18-FEB-2000; 2000WO-US004341.  
PR 18-FEB-2000; 2000WO-US004342.  
PR 24-FEB-2000; 2000WO-US004414.  
PR 24-FEB-2000; 2000WO-US004914.  
PR 24-FEB-2000; 2000WO-US005004.  
PR 01-MAR-2000; 2000WO-US005601.  
PR 02-MAR-2000; 2000WO-US005746.  
PR 02-MAR-2000; 2000WO-US005841.  
PR 15-MAR-2000; 2000WO-US006319.  
PR 15-MAR-2000; 2000WO-US006884.  
PR 20-MAR-2000; 2000WO-US007377.  
PR 21-MAR-2000; 2000WO-US007532.  
PR 30-MAR-2000; 2000WO-US008439.  
PR 17-MAY-2000; 2000WO-US013705.  
PR 22-MAY-2000; 2000WO-US014042.  
PR 30-MAY-2000; 2000WO-US014941.  
PR 02-JUN-2000; 2000WO-US015264.  
PR 28-JUL-2000; 2000WO-US020710.  
PR 11-AUG-2000; 2000WO-US022031.  
PR 23-AUG-2000; 2000WO-US023522.  
PR 24-AUG-2000; 2000WO-US023328.  
PR 08-NOV-2000; 2000WO-US030952.  
PR 10-NOV-2000; 2000WO-US030873.  
PR 01-DEC-2000; 2000WO-US032678.  
PR 20-DEC-2000; 2000US-00747259.



```

PR 20-DEC-2000; 2000WO-US034956.
PR 28-FEB-2001; 2001US-00796498.
PR 28-FEB-2001; 2001WO-US006520.
PR 01-MAR-2001; 2001WO-US006666.
PR 09-MAR-2001; 2001US-00802706.
PR 14-MAR-2001; 2001US-00808689.
PR 22-MAR-2001; 2001US-00816744.
PR 05-APR-2001; 2001US-00828366.
PR 10-MAY-2001; 2001US-00854208.
PR 10-MAY-2001; 2001US-00854280.
PR 18-MAY-2001; 2001US-00860216.
PR 25-MAY-2001; 2001US-00866028.
PR 25-MAY-2001; 2001US-00866034.
PR 01-JUN-2001; 2001WO-US017092.
PR 01-JUN-2001; 2001US-00872035.
PR 01-JUN-2001; 2001WO-US017800.
PR 05-JUN-2001; 2001US-00874503.
PR 14-JUN-2001; 2001US-00882636.
PR 19-JUN-2001; 2001US-00886342.
PR 20-JUN-2001; 2001WO-US019692.
PR 21-JUN-2001; 2001US-00887879.
PR 22-JUN-2001; 2001WO-US020116.
PR 29-JUN-2001; 2001WO-US021056.
PR 09-JUL-2001; 2001WO-US021735.
PR 18-JUL-2001; 2001US-00908827.
PR 06-AUG-2001; 2001US-00924419.
PR 09-AUG-2001; 2001US-00927796.
PR 16-AUG-2001; 2001US-00931836.
PR 19-DEC-2001; 2001US-00028072.
XX
XX (GETH ) GENENTECH INC.
XX
PI Baker KP, Beresini M, Deforge L, Desnoyers L, Filvaroff E, Gao W;
PI Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;
PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;
XX
DR WPI; 2003-695927/66.
DR N-PSDB; ADB19071.
XX
XX Novel secreted and transmembrane PRO polypeptides useful for stimulating
XX the release of tumor necrosis factor alpha and detecting the presence of
XX a tumor in a mammal.
XX
XX Claim 12; Fig 64; 660pp; English.
XX
XX The invention describes 305 nucleic acids encoding PRO (secreted and
XX transmembrane) polypeptides (I). (I) is useful for stimulating the
XX release of TNF-alpha from human blood, for modulating the uptake of
XX glucose or FFA by skeletal muscle cells or adipocyt
XX
XX Sequence 504 AA;
XX
Query Match 100.0%; Score 2615; DB 6; Length 504;
Best Local Similarity 100.0%; Pred. No. 1.8e-258;
Matches 504; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEATGTWALLALLALLLLTALSCTRARGHLPPGTPPLGNNLQLRPGALYSGLMRL 60
DB 1 MEATGTWALLALLALLLLTALSCTRARGHLPPGTPPLGNNLQLRPGALYSGLMRL 60

QY 61 SKYGPVFTIYLGPRPVPVVVVGQEAVERALGQAEFEFGSGRTVAMLEGTFDGHGVFFSN 120
DB 61 SKYGPVFTIYLGPRPVPVVVVGQEAVERALGQAEFEFGSGRTVAMLEGTFDGHGVFFSN 120

QY 121 GERWRQLRKFTMLALDLGNKRGEBELIQAEARCLVETFGTEGPRFPDPSLLAQATSN 180
DB 121 GERWRQLRKFTMLALDLGNKRGEBELIQAEARCLVETFGTEGPRFPDPSLLAQATSN 180

QY 181 VVCSLLFGLRFSEYDEKFEQAVVRAAGTLLGVSSQGGQTYEMFSPWLRPLPGPHKOLLHH 240
DB 181 VVCSLLFGLRFSEYDEKFEQAVVRAAGTLLGVSSQGGQTYEMFSPWLRPLPGPHKOLLHH 240

QY 241 VSTLAFTVRQVQOQHGNLDASGPARDLVDAFLKMAQEQNPGEFTNKNMLMTVIYLL 300

```

Search completed: September 15, 2004, 09:18:33  
Job time : 134 secs

```

DB 241 VSTLAFTVRQVQOQHGNLDASGPARDLVDAFLKMAQEQNPGEFTNKNMLMTVIYLL 300
QY 301 FAGTMTVSTTVGYTLLLLMKYPHVOKWVREELNRELGAGQAPSLGDRTRLPYTDAVHEA 360
DB 301 FAGTMTVSTTVGYTLLLLMKYPHVOKWVREELNRELGAGQAPSLGDRTRLPYTDAVHEA 360
QY 361 QRLALVPMGIPRTLMRTTRFRGYTLPGQTEVFPLLSILHDPNIFKHPPEENPDRFLDA 420
DB 361 QRLALVPMGIPRTLMRTTRFRGYTLPGQTEVFPLLSILHDPNIFKHPPEENPDRFLDA 420
QY 421 DGRFRKHEAFLPFSLGKRVCLGEGIAKAELELFPFTTILQAFSLSPCPDPTLSLKPTVSG 480
DB 421 DGRFRKHEAFLPFSLGKRVCLGEGIAKAELELFPFTTILQAFSLSPCPDPTLSLKPTVSG 480
QY 481 LFNIPPAFQLOVRPTDLHSTTTQTR 504
DB 481 LFNIPPAFQLOVRPTDLHSTTTQTR 504

```

***This Page Blank (uspto)***